

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.

THIS PAGE BLANK (USPTO)

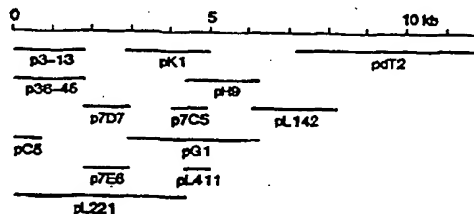


INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁵: C12N 15/40, 15/82, 5/10, A01H 5/00	A2	(11) International Publication Number: WO 94/21796 (43) International Publication Date: 29 September 1994 (29.09.94)
(21) International Application Number: PCT/US94/03028 (22) International Filing Date: 22 March 1994 (22.03.94) (30) Priority Data: 08/038,768 24 March 1993 (24.03.93) US (71) Applicants: PIONEER HI-BRED INTERNATIONAL, INC. [US/US]; 700 Capital Square, 400 Locust Street, Des Moines, IA 50309 (US). THE UNITED STATES OF AMERICA as represented by THE SECRETARY, UNITED STATES DEPARTMENT OF AGRICULTURE [US/US]; 12th and Independence Avenue, S.W., Washington, DC 20250-1400 (US). (72) Inventors: ROTH, Bradley, A.; 210 N.W. 3rd Place, Grimes, IA 50111 (US). TOWNSEND, Rod; 541 Waterbury Circle, Des Moines, IA 50312 (US). MCMULLEN, Michael, D.; 1680 Madison Avenue, Wooster, OH 44691-4096 (US). (74) Agents: ROTH, Michael, J. et al.; 700 Capital Square, 400 Locust Street, Des Moines, IA 50309 (US).		(81) Designated States: AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, ES, FI, GB, HU, JP, KP, KR, KZ, LK, LU, LV, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SK, UA, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). Published <i>Without international search report and to be republished upon receipt of that report.</i>

(54) Title: MAIZE CHLOROTIC DWARF VIRUS AND RESISTANCE THERETO**(57) Abstract**

Methods and materials are provided to isolate the coat protein genes from maize chlorotic dwarf virus. One or more of these genes (MCDV-CP₁, MCDV-CP₂ or MCDV-CP₃) is then incorporated in an expression cassette designed for suitable expression in a plant cell system. The resulting transformation vector is then introduced into maize to provide cross protection to MCDV or related viral infections.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

MAIZE CHLOROTIC DWARF VIRUS AND RESISTANCE THERETO

5

Technical Field

This invention relates to providing plants with resistance to maize chlorotic dwarf virus (MCDV) and viruses to which MCDV infection or resistance provides cross-resistance, including maize dwarf mosaic virus strain A (MDMV-A).

Background of the Invention

Virus-induced diseases in agronomically important crops have cost farmers a great loss of income due to reduced yields. Traditionally, virus diseases have been controlled by breeding for host plant resistance or by controlling insects that transmit diseases. Chemical means of protection are not generally possible for most viruses, and where possible are not generally practical. It has been known for many years that viral symptoms can be reduced in virus-infected plants by prior inoculation with a mild strain of the same virus, a phenomena known as cross-protection, as described by Sequeira, L., Trends in Biotechnology, 2, 25 (1984). Cross-protection is considered successful if the disease symptoms of the superinfecting (the more virulent) virus can be delayed or suppressed. There are several disadvantages to applying this type of cross-protection to the field situation:

- 1) application of the mild strain virus to entire fields is usually not practical,
- 2) the mild strain might undergo mutation to a more highly virulent strain,
- 3) the protecting strain might interact synergistically with a non-related virus causing a severe pathogenic infection,
- 4) a protecting virus in one crop may be a severe pathogen in another crop, and
- 5) a protective strain may cause a significant loss of yield in itself.

One proposed solution to these disadvantages has been to introduce a single viral gene into the host plant genome to cross-protect, rather than infect with an intact virus. This single gene cross-protection strategy has already been proven successful using the coat protein gene from tobacco mosaic virus (TMV-CP). As

reported by Abel, P.P., et.al., Science, 232, 738 (1986), transgenic tobacco plants, expressing TMV mRNA and coat protein (CP), demonstrated delayed or suppressed symptom development upon infection with TMV. TMV-CP transgenic tomato plants have been described by Nelson, R.S., et.al., Bio/Technology, 6, 403 (1988), to show evidence of protection from TMV as well as three strains of tomato mosaic virus (ToMV). Other approaches using DNA clones of viruses to engineer resistance include positive interference, as described by Golemboski et al. Proc. Natl. Acad. Sci. USA, 87, 6311 (1990) and Carr and Zaitlin, Mol. Pl. Microbe Inter., 4, 579 (1991); and antisense RNA, as described by Powell et al., Proc. Natl. Acad. Sci. USA, 86, 6949 (1989).

Numerous viruses exist for which resistance is desired. Maize chlorotic dwarf virus causes a somewhat variable mosaic or yellow streaking and occasional stunting in maize. Early infections can result in severe symptoms including premature death. The virus is spread by the blackfaced leafhopper (*Graminella nigrifrons*). MCDV can overwinter in Johnsongrass (*Sorghum halepense*) and as a result has become a recurrent problem in areas where Johnsongrass is a common weed. Combined infections with maize dwarf mosaic virus can cause more severe symptoms although the syndrome is less well characterized than Corn Lethal Necrosis. Only limited success has been obtained to date in developing MCDV-resistant maize lines, due to the difficulties of selecting efficiently for resistance to an obligately insect transmitted virus, as well as a lack of usable sources of resistance in agronomically useful maize lines. Thus, there is a continuing need for genes, plant transformation vectors, and transformed plant materials providing resistances to pathogenic viruses such as MCDV.

Unfortunately, while certain plant viruses, such as tobacco mosaic virus, have coat protein genes that are found on subgenomic RNA and are therefore relatively easy to identify and clone for use in engineered cross-protection, maize chlorotic dwarf virus belongs to a completely separate group, the only other (tentatively assigned) member of which is the spherical virus of the rice tungro disease (RTSV). In addition, MCDV has a number of unusual biological properties which make identification of an appropriate gene difficult. For example, all attempts to mechanically transmit MCDV have been unsuccessful. As another example, MCDV appears to be a phloem-restricted virus. MCDV also has three coat proteins, and it was not known whether expression of one protein would be sufficient to confer immunity or whether all three would need to be expressed. Nor was it known

which protein would be the appropriate one to express if only one could be expressed. Further, the genome of MCDV has an unusual genome organization to provide for the expression of multiple coat proteins.

Brief Description of the Drawing Figures

5 Figure 1 is a schematic illustration of the manner in which the nucleic acid sequence of MCDV-type strain was obtained by sequencing overlapping cDNA clones.

Figure 2 is an a schematic illustration of the unusual organization of the MCDV genome.

10 Disclosure of the Invention

In the present invention, methods and materials are provided to isolate any or all of the three coat protein genes from maize chlorotic dwarf virus (MCDV). One or more of these genes (MCDV-CP_x, where x is 1, 2, or 3) is then incorporated in an expression cassette designed for suitable expression in a plant cell system.

15 The resulting transformation vector is then introduced into maize callus to provide cross-protection to MCDV-related viral infections. MCDV has a single, long RNA core having the sequence shown in SEQUENCE I.D. No. 4.

Description of the Preferred Embodiments

The present invention provides cDNA clones from the RNA genome of maize chlorotic dwarf virus which code substantially solely for the coat protein of the virus. These clones are incorporated into an expression cassette in which the cDNA clone is operably linked to plant or bacterial regulatory sequences which cause the expression of the cDNA clone in living plant or bacterial cells, respectively. It is important that the cloned gene have a start codon in the correct reading frame for the structural sequence. The resulting bacterial vectors can be readily inserted into

20 bacteria for expression and characterization of the sequence. Accordingly, the present invention also provides bacterial cells containing as a foreign plasmid at least one copy of the foregoing bacterial expression cassette. In addition, the plant expression cassette preferably includes a strong constitutive promoter sequence at one end to cause the gene to be transcribed at a high level and a poly-A recognition sequence at the other end for proper processing and transport of the messenger RNA. An example of such a preferred (empty) expression cassette into which the cDNA of the present invention can be inserted is the pPHI414 plasmid developed by Beach et al. of Pioneer Hi-Bred International, Inc., Johnston, IA, as disclosed in

25 30 35 U.S. Patent Application No. 07/785,648, filed October 31, 1991. Highly preferred

plant expression cassettes will be designed to include one or more selectable marker genes, such as kanamycin resistance or herbicide tolerance genes. The plant expression vectors of this invention can be inserted, using any convenient technique, including electroporation (in protoplasts), microprojectile bombardment, and microinjection, into cells from monocotyledonous or dicotyledonous plants, in cell or tissue culture, to provide transformed plant cells containing as foreign DNA at least one copy of the DNA sequence of the plant expression cassette. Preferably, the monocotyledonous species will be selected from maize, sorghum, wheat and rice, and the dicotyledonous species will be selected from soybean, alfalfa, tobacco and tomato. Using known techniques, protoplasts can be regenerated and cell or tissue culture can be regenerated to form whole fertile plants which carry and express the desired cDNA clone for MCDV coat protein. Accordingly, a highly preferred embodiment of the present invention is a transformed maize plant, the cells of which contain as foreign DNA at least one copy of the DNA sequence of an expression cassette of this invention.

Finally, this invention provides methods of imparting resistance to maize chlorotic dwarf virus to plants of a MCDV susceptible taxon, comprising the steps of:

- a) culturing cells or tissues from at least one plant from the taxon,
- b) introducing into the cells of the cell culture or tissue culture at least one copy of an expression cassette comprising a cDNA clone from the RNA genome of MCDV which codes substantially solely for the coat protein of the virus, operably linked to plant regulatory sequences which cause the expression of the cDNA clone in the cells, and
- c) regenerating MCDV-resistant whole plants from the cell or tissue culture. Once whole plants have been obtained, they can be sexually or clonally reproduced in such manner that at least one copy of the sequence provided by the expression cassette is present in the cells of progeny of the reproduction.

Alternatively, once a single transformed plant has been obtained by the foregoing recombinant DNA method, conventional plant breeding methods can be used to transfer the coat protein gene and associated regulatory sequence via crossing and backcrossing. Such intermediate methods will comprise the further steps of

- a) sexually crossing the MCDV resistant plant with a plant from the MCDV susceptible taxon;

- b) recovering reproductive material from the progeny of the cross; and
 - c) growing resistant plants from the reproductive material. Where desirable or necessary, the characteristics of the susceptible taxon can be substantially preserved by expanding this method to include the further steps of repetitively:
 - a) backcrossing the MCDV resistant progeny with MCDV susceptible plants from the susceptible taxon; and
 - b) selecting for expression of MCDV resistance among the progeny of the backcross,
- until the desired percentage of the characteristics of the susceptible taxon are present in the progeny along with the gene imparting MCDV resistance.

By the term "taxon" herein is meant a unit of botanical classification of genus or lower. It thus includes genus, species, cultivars, varieties, variants, and other minor taxonomic groups which lack a consistent nomenclature.

- It will also be appreciated by those of ordinary skill that the plant vectors provided herein can be incorporated into Agrobacterium tumefaciens or Agrobacterium rhizogenes, which can then be used to transfer the vector into susceptible plant cells, primarily from dicotyledonous species. Thus, this invention provides a method for imparting MCDV resistance in Agrobacterium-susceptible dicotyledonous plants in which the expression cassette is introduced into the cells by infecting the cells with Agrobacterium tumefaciens, a plasmid of which has been modified to include the plant expression cassette of this invention. The following description further exemplifies the compositions of this invention and the methods of making and using them. However, it will be understood that other methods, known by those of ordinary skill in the art to be equivalent, can also be employed.

1. Isolation and cloning of MCDV cDNA

- The type strain of MCDV was maintained in the maize inbred Oh28 by transmission with the leafhopper G. nigrifrons and viral particles were isolated as previously described (Hunt et al., Phytopathology 78, 449 (1988)). MCDV particles were suspended in NETS (10 mM Tris, pH 7.5; 100 mM NaCl; 1 mM Na₂EDTA; 0.5% SDS) and extracted with 1:1 chloroform:phenol to isolate MCDV RNA.

- First and second strand cDNA synthesis were by the method of Gubler and Hoffman, Gene 25, 263 (1983) utilizing cDNA synthesis kits (Amersham, Arlington Heights, IL). For the initial cDNA libraries, double-stranded cDNA was treated

with EcoRI methylase, ligated to GGAATTCC EcoRI linkers, digested with EcoRI and separated from linkers by column fractionation. The cDNA was ligated to EcoRI-cleaved _gt10 and EcoRI-cleaved, phosphatased (CIP) _gt11 phage arms. After packaging, the _gt10 phage were plated on bacterial strain NM514 and
5 screened for MCDV-specific inserts by filter plaque hybridization (Benton and Davis, Science 196, 180 (1977)), using ^{32}P -labeled cDNA's random-primed from the MCDV genomic RNA. MCDV-positive phage were purified and the cDNA inserts subcloned into pUC119 (Vieira and Messing, Meth. Enzymol. 153, 3 (1987)) for further analysis. Hybridization positive clones from the initial _gt10 library
10 included: p3-13, p36-45, pH9, pK1, pG1, pC5 (Figure 1). After packaging, the _gt11 phage were plated on bacterial strain Y1090^r and screened with antisera to either intact MCDV virions or isolated, individual MCDV capsid proteins (Maroon, MS Thesis, Ohio State University (1989)) as described by Mierendorf, et al. Meth. Enzymol. 152,458 (1987). Positive phage clones were identified with antisera
15 specific to either cp1 or cp2, and cDNA inserts from these phage were subcloned into pUC119. The anti-cp1-specific cDNA clone, p7C5, and the anti-cp2-specific cDNA clones, p7E6 and p7D7, (Figure 1) were chosen for study. Analysis of initial cDNAs revealed that a number of clones terminated at identical EcoRI sites which were shown to be present in the viral sequence. This result indicated that the
20 methylation of the initial cDNAs was incomplete. To obtain cDNAs to the rest of MCDV and to overlap the initial clones, two additional cDNA libraries were prepared, one primed with oligo-dT(12-18) and one random-primed. Double-stranded cDNA prepared as above was ligated to a 20/24 nt. blunt end/EcoRI adaptor (Amersham), and adaptor cDNAs were kinased and ligated to
25 EcoRI-cleaved/phosphatased pUC119. Plasmid clone pdT2 (Figure 1) was derived from the dT-primed library and plasmids pL142, pL221, and pL411 (Figure 1) were derived from the random-primed library.

2. Sequencing of MCDV cDNA

Single-stranded DNA templates for sequencing were derived by
30 superinfection with M13K07 of bacterial strain MV1190 containing the pUC119 based cDNAs (Figure 1), cloned in both orientations, as described by McMullen et al., Nuc. Acids Res. 14, 4953 (1986) and Vieira and Messing, Meth. Enzymol. 153, 3 (1987). Ordered deletions from the full-length single-stranded templates were prepared by the method of Dale et al., Plasmid 13, 31 (1985). Dideoxynucleotide
35 sequencing reactions with the Klenow fragment of Pol. I or Sequenase (U.S.

Biochemicals, Cleveland, OH) were performed using ^{35}S -dATP. Greater than 99% of the total sequence was obtained from both strands and the majority was read from three or more templates. The 5' sequence not contained on cDNA was obtained by direct RNA sequencing, using the sequencing primer 5'-GGTCTACTCACGGCAGCCA-3' (SEQUENCE I.D. NO. 3) with an RNA sequencing kit (Boehringer Mannheim, Indianapolis, IN) as recommended except that tailing of reaction products with dTTP by terminal deoxynucleotidyl transferase using the method of DeBorde *et al.*, Anal. Biochem. 157,275 (1986) was added to improve resolution of final bases.

To obtain the amino-terminal protein sequence of MCDV capsid proteins, MCDV particles were disrupted in Laemmli loading buffer and the individual capsid protein separated on a 12.5%-4% Laemmli slab gel (Laemmli, Nature 227, 680 (1970)). The proteins were electrotransferred to Immobilon-P membrane (Millipore, Bedford, MA) using a 10 mM CAPS, pH 11.0; 10% MeOH transfer buffer, stained with Coomassie Blue R-250 for visualization and excised. Automated amino-terminal protein sequencing was performed by the Iowa State University Biochemistry Instrumentation Center (Ames, IA).

DNA and protein sequence analysis was performed using the IntelliGenetics (Mountain View, CA) molecular biology software on a Digital VAX 8250 located at the USDA-ARS-ASRR (Agricultural Systems Research Resource) Beltsville, MD.

The nucleic acid sequence of MCDV-type strain was obtained by sequencing overlapping cDNA clones (Figure 1) that covered all but 13 nucleotides at the 5' terminus of MCDV. The 5' end sequence was obtained by direct RNA sequencing. Despite repeated attempts and the use of terminal transferase in the manner of DeBorde *et al.*, Anal. Biochem. 157, 275 (1986) the first nucleotide could not be definitely determined. In part for this reason, the expressions "coding substantially for" and "coding substantially solely for" are used herein, and with regard to the use of the word "substantially" refer to sequences which code for no more than a few (five or less) amino acids greater or lesser on either end of the desired protein or proteins, or which have an equivalent number of nucleotide-bases more or less than the native sequence.

The genomic RNA of MCDV-type (SEQUENCE I.D. NO. 4) was determined to be 11785 nucleotides long, exclusive of the poly-A tail at the 3' terminus. This sequence permits the construction of a DNA molecule which codes for the entire maize chlorotic dwarf virus, or any portion or functional unit thereof which is

useful in conferring resistance to the virus when expressed in plant cells. Such resistance can readily be evaluated using routine testing methods such as those disclosed herein. Computer analysis of the sequence indicated a long open reading frame from nucleotide 456 to nucleotide 10826. The translation of this open reading frame would result in a protein of 3457 amino acids with a derived molecular weight of 388,890 daltons. The open reading frame begins with two AUG triplets, neither of which is in a particularly favorable context for initiation of translation when compared with the analyses of translation start sequences by Lutcke *et al.*, EMBO J. 6, 43 (1987); and Kosak, J. Cell Biol. 108, 229 (1989) by the scanning model. In addition, there are 13 AUG triplets preceding the double AUG that starts the open reading frame. A long untranslated 3' leader containing multiple AUG triplets before the beginning of a very long reading frame is similar to the animal picornaviruses as described by Stanway, J. Gen. Virol. 71, 2483 (1990). Internal initiation at the AUG for the long open reading frame has been demonstrated to occur for a number of the animal picornaviruses as seen in Pelletier and Sonenberg, Nature, 334, 320 (1988) and Jang *et al.*, J. Virol. 63, 1651 (1989). The mechanism for initiation of translation for MCDV has not been characterized.

The derived amino acid sequence of MCDV-type was compared to the Protein Identification Resource, Version 32 and the University of Geneva, Version 22, protein data banks for sequence similarity using the IFIND (IntelliGenetics) program based on the algorithm of Wilber and Lipman, Proc. Natl. Acad. Sci. USA, 80, 726 (1983). The highest similarity score was with the comovirus, cowpea mosaic virus (CPMV) as reported by Lomonosoff and Shanks, EMBO J. 2, 2253 (1983) and the second highest score was with the nepovirus, grapevine fanleaf virus (GFLV) as reported by Ritzenthaler *et al.*, J. Gen. Virol. 72, 2357 (1991). For both viruses the region of similarity preceded and included the first conserved motif of RNA-dependent RNA polymerases as defined by Poch *et al.* EMBO J. 8, 3867 (1989). The IFIND program identified weaker similarity with additional nepoviruses and some of the animal picornaviruses. The conservation of protein sequence and gene order for the plant comoviruses, nepoviruses and potyviruses, and the animal picornaviruses is well documented by, *inter alia*, Agros *et al.*, Nuc. Acids Res. 12, 7251 (1984); Goldbach, Ann. Rev. Phytopath. 24, 289 (1986); and Domier *et al.*, Virology, 158, 20 (1987) and has led to the proposal of the picornavirus-like "supergroup". Two additional conserved protein regions involved in genome replication for picorna-like viruses are the NTP binding/helicase region,

as described by Agros *et al.*, above, and Gorbalenya *et al.*, *Nuc. Acids Res.*, 17, 4713 (1989) and the C-terminal region, cysteine active site of the 3C-like proteases, as also described by Agros *et al.*, above, and by Grief *et al.*, *J. Gen. Virol.*, 69, 1517 (1988).

The electrophoresis of MCDV virions on denaturing protein gels reveals
5 three structural proteins, designated cp1, cp2 and cp3 with molecular weights of 32.5 kd, 27 kd, and 24.5 kd; respectively. Antiserum specific to cp1 was used to screen a *gt11* library to isolate the clone p7C5, and antiserum specific to cp2 was used to identify the cDNAs p7E6 and p7D7 (Figure 1). This result indicated that an antigenic region of cp1 was located between 4063-4903 and an antigenic region of
10 cp2 was located between 1815-2941. Automated amino-terminal sequencing was performed on each of the MCDV capsid proteins. The amino-terminus of cp2 was apparently blocked as no sequence was obtained. The 15 amino acids at the NH₂-terminus of cp3 were determined to be LQVASLTDIGELSSV, as shown in SEQUENCE I.D. NO. 2 and SEQUENCE I.D. NO. 6. This sequence is an exact
15 match to the derived protein sequence encoded by nucleotides 3144-3188. Likewise, the 15 amino acids at the NH₂-terminus of cp1, VSLGRSFENGVLIGS, as shown in SEQUENCE I.D. NO. 5 and SEQUENCE I.D. NO. 7, are an exact match to the derived protein sequence encoded by nucleotides 3750-3794. Both proteins must be derived by proteolytic cleavage of the large polyprotein. The Gln/Leu cleavage at
20 the NH₂-terminus of cp3 and Gln/Val cleavage at the NH₂-terminus of cp1 are dipeptide cleavage sites that may be used by animal picornavirus 3C proteases, according to Krausslich and Wimmer, *Ann. Rev. Biochem.*, 57, 754 (1988), which could indicate that the 3C-similar region of the MCDV may function in capsid protein processing. Assuming that cp3 begins with the Leu at the Gln/Leu cleavage and ends with the Gln at the Gln/Val cleavage for cp1, cp3 would have a derived
25 MW of 21,933, a little less than the 24.5 kd MW determined by SDS gel electrophoresis. Although protein sequence was not obtained for cp2, the position of clones p7E6 and p7D7, and the finding that protein fusions expressed from the pEX vector for the *Pst*I fragments 2076-2619 and 2613-3149 reacted positively with
30 cp2-specific antiserum (McMullen, unpublished), is consistent with cp2 preceding cp3 in the polyprotein similar to the order of vp2-vp3-vp1 for the animal picornaviruses. However, it is still not known if the coding region for cp2 immediately precedes cp3.

The overall genome structure of MCDV-type strain is shown in Figure 2.
35 MCDV genome organization resembled that of the animal picornaviruses, a single

large polyprotein in which the capsid proteins are encoded 5' of the proteins presumed to be involved in genome replication. Depending on the exact location of cp2, the MCDV genome can encode up to 78 kd of protein 5' of the capsid proteins for which there are no corresponding animal picornavirus protein. This region may encode plant virus specific functions such as cell-to-cell movement or helper protein for insect transmission. Because MCDV is a phloem restricted virus, there is no evidence for a virus-encoded cell-to-cell movement protein. However, there is evidence for the presence of an insect transmission helper component in MCDV-infected plants according to Hunt *et al.*, Phytopathology, 78, 449 (1988). The presence of plant-virus-specific proteins at the NH₂-terminus of the polyprotein would allow addition of these proteins without disruption of the cp proteins-replication functions genome structure typical of picornaviruses.

3. Design of the plasmid vector.

The gene MCDV coat protein 3 was placed under control of tandem cauliflower mosaic virus 35S promoters isolated from the 1841 strain of the virus, and a polyadenylation signal sequence obtained from the potato proteinase inhibitor II (Pin II) gene that exhibits enhancer-like activity. The chimeric gene also included a 79 bp sequence Ω' from the 5' leader region of tobacco mosaic virus (TMV) that functions as a translational enhancer; and a Zea mays alcohol dehydrogenase 1, intron 1 fragment (ADH) spanning nucleotides 119-672, trimmed to 557 bp with Bal 31 nuclease, which has been shown to function as an enhancer of gene expression in monocots. The plasmids were grown in *E. coli* and purified by the known polyethylene glycol precipitation method of Sambrook *et al.*, Molecular Cloning, 1, 40 (1989). Purity was confirmed by electrophoretic analysis of the DNA fragments obtained after digestion with restriction endonucleases. The plasmid was designated pPHI1406 and the sequence is shown in SEQUENCE I.D. No. 1.

4. Preparation of the recipient organism.

Separately, an embryogenic cell suspension line 54-68-5 was established from immature embryos obtained from a cross between a line derived from the public inbred corn line B73 and a WX 1-9 translocation stock of public inbred corn line W23.

5. Transformation

Suspension cells from (4) were bombarded with 1 μ l aliquots of a 30 μ l mixture containing 10 μ g of purified plasmid DNA (5 μ g of the MCDV plasmid pPHI1406 (SEQUENCE I.D. No. 1), and 5 μ g of the same plasmid in which the BAR

(Basta resistance) gene was substituted for the MCDV cp3 gene) precipitated onto 1 μ m tungsten particles as described by numerous articles including Klein, T.M., et al., 1988 (May) Bio/Technology 6:559-563; Klein, T. M., et al., 1988 (June) Proc. Natl. Acad. Sci. USA 85:4305-4309; T. M. Klein, et al., "Stable Genetic Transformation of Intact Applicant Nicotiana Cells by the Particle Bombardment Process", Proc. Natl. Acad. Sci. USA, Vol. 85, November 1988, pp. 8502-8505; D. T. Tomes, et al., "Transgenic Tobacco Plants and their Progeny Derived by Microprojectile Bombardment of Tobacco Leaves", Plant Molecular Biology, Vol. 14, No. 2, February, 1990, pp. 261-268, Kluwer Academic Publishers, BE; and M. C. Ross, et al., "Transient and Stable Transgenic Cells and Calli of Tobacco and Maize Following Microprojectile Bombardment", J. Cell. Biochem., Suppl. 13D, 27th March - April 1989, P. 268, Abstract No. M. 149, Alan R. Liss, Inc. New York, US; and plated onto selective medium containing 5 ppb phosphinothricin (Basta™).

Following a prolonged period of selection and callus growth, regeneration was initiated by placing callus on a Murashige & Skoog medium modified by addition of 0.5 mg/l 2,4-D and 5 ppb Basta. Embryogenic callus was selected and transferred to medium lacking 2,4-D and kept in a lighted growth room. Germinated plantlets were placed in culture tubes and finally planted out into soil in pots in the greenhouse.

More than 150 R₀ (recombinant) plants were obtained, representing twenty independent transformation events. Transformation was confirmed by PCR amplification of a DNA fragment spanning part of the MCDV coat protein gene and the CaMV promoter. Genomic DNA samples, in which a fragment of the expected size was successfully amplified were presumed to be transformed. These plants were pollinated with pollen from non-transgenic B73 plants and the resulting R₁ seed was planted in a field trial under USDA supervision. The resulting plants exhibited a virus resistant phenotype, i.e., they survived and set seed under virus infection conditions in which non-transgenic plants died prematurely, as seen in the following table:

Field Test Results

	<u>Transgenic*</u>	<u>Control</u>
Number of Plants	379	32
Number of Harvestable Ears	52	0
% Harvested vs. Total	13.7%	0%

The screening was performed in a manner to insure maximum infection levels and severity. Thus, the level of resistance seen in this extreme test corresponds to effective, usable virus tolerance when the transformants of this invention are used under normal farming conditions.

- 5 The MCDV resistance is a simply inherited, dominant trait and can, if desired, be introduced into other maize varieties by simple crossing or backcrossing. In addition to providing resistance to MCDV, this invention is also capable of conferring resistance to viruses to which plants obtain cross-resistance through infection by MCDV. In the field test described above, resistance to maize
10 dwarf mosaic virus strain A (MDMV-A) was also observed. Accordingly, this invention provides resistance to that virus as well.

SEQUENCE LISTING

- (1) GENERAL INFORMATION:
- (i) APPLICANT: McMullen, Michael D.; Roth, Bradley A.; Townsend, Rod
- 5 (ii) TITLE OF INVENTION: MAIZE CHLOROTIC DWARF VIRUS RESISTANCE
- (iii) NUMBER OF SEQUENCES: 7
- (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: Pioneer Hi-Bred International, Inc.
- (B) STREET: 700 Capital Square, 400 Locust
- 10 Street
- (C) CITY: Des Moines
- (D) STATE: Iowa
- (E) COUNTRY: United States
- (F) ZIP: 50309
- 15 (v) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
- (B) COMPUTER: IBM Compatible
- (C) OPERATING SYSTEM: MS-DOS, Microsoft Windows
- (D) SOFTWARE: Microsoft Windows Notepad
- 20 (vi) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER:
- (B) FILING DATE:
- (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
- 25 (A) APPLICATION NUMBER:
- (B) FILING DATE:
- (viii) ATTORNEY/AGENT INFORMATION:
- (A) NAME: Roth, Michael J.
- (B) REGISTRATION NUMBER: 29,342
- 30 (C) REFERENCE/DOCKET NUMBER: 0235 US
- (ix) TELECOMMUNICATION INFORMATION:
- (A) TELEPHONE: (515) 245-3594
- (B) TELEFAX: (515) 245-3634
- (2) INFORMATION FOR SEQ ID NO: 1:
- 35 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5033 base pairs
 (B) TYPE: nucleotide
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear
- 5 (ii) MOLECULE TYPE: synthetic DNA
 (A) DESCRIPTION: transformation plasmid pPHI1406
 (iii) HYPOTHETICAL: No
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
- | | | | | | | |
|----|------------|------------|------------|------------|-------------|------|
| | TCGCGCGTTT | CGGTGATGAC | GGTGAAAACC | TCTGACACAT | GCAGCTCCCC | 50 |
| 10 | GAGACGGTCA | CAGCTTGTCT | GTAAGCGGAT | GCCGGGAGCA | GACAAGCCCC | 100 |
| | TCAGGGCGCG | TCAGCGGGTG | TTGGCGGGTG | TCGGGGCTGG | CTTAACATATG | 150 |
| | CGGCATCAGA | GCAGATTGTA | CTGAGAGTGC | ACCATATGCG | GTGTGAAATA | 200 |
| | CCGCACAGAT | GCGTAAGGAG | AAAATACCGC | ATCAGGCGCC | ATTCGCCATT | 250 |
| | CAGGCTGCGC | AACTGTTGGG | AAGGGCGATC | GGTGCGGGCC | TCTTCGCTAT | 300 |
| 15 | TACGCCAGCT | GGCGAAAGGG | GGATGTGCTG | CAAGGCGATT | AAGTTGGGTA | 350 |
| | ACGCCAGGGT | TTTCCAGTCT | ACGACGTTGT | AAAACGACGG | CCAGTGCCAA | 400 |
| | GCTCAGATCT | GAGCTTCTAG | AAATCCGTCA | ACATGGTGA | GCACGACACT | 450 |
| | CTCGTCTACT | CCAAGAATAT | CAAAGATACA | GTCTCAGAAG | ACCAAAGGGC | 500 |
| | TATTGAGACT | TTTCAACAAA | GGGTAATATC | GGGAAACCTC | CTCGGATTCC | 550 |
| 20 | ATTGCCCAGC | TATCTGTCA | TTCATCAAAA | GGACAGTAGA | AAAGGAAGGT | 600 |
| | GGCACCTACA | AATGCCATCA | TTGCGATAAA | GGAAAGGCTA | TCGTTCAAGA | 650 |
| | TGCCTCTGCC | GACAGTGGTC | CCAAAGATGG | ACCCCCACCC | ACGAGGAGCA | 700 |
| | TCGTGGAAAA | AGAAGACGTT | CCAACCACGT | CTTCAAAGCA | AGTGGATTGA | 750 |
| | TGTGATGCTC | TAGAAATCCG | TCAACATGGT | GGAGCACGAC | ACTCTCGTCT | 800 |
| 25 | ACTCCAAGAA | TATCAAAGAT | ACAGTCTCAG | AAGACCAAAG | GGCTATTGAG | 850 |
| | ACTTTTCAAC | AAAGGGTAAT | ATCGGGAAAC | CTCCTCGGAT | TCCATTGCCC | 900 |
| | AGCTATCTGT | CACTTCATCA | AAAGGACAGT | AGAAAAGGAA | GGTGGCACCT | 950 |
| | ACAAATGCCA | TCATTGCGAT | AAAGGAAAGG | CTATCGTTCA | AGATGCCTCT | 1000 |
| | GCCGACAGTG | GTCCCAAAGA | TGGACCCCCA | CCCACGAGGA | GCATCGTGGA | 1050 |
| 30 | AAAAGAAGAC | GTTCCAACCA | CGTCTTCAAA | GCAAGTGGAT | TGATGTGATA | 1100 |
| | TCTCCACTGA | CGTAAGGGAT | GACGCACAA | CCCCTATCC | TTGCAAGAC | 1150 |
| | CCTTCCTCTA | TATAAGGAAG | TTCATTTCAT | TTGGAGAGGA | CGAGCTGCAG | 1200 |
| | CTTATTTT | CAACAATTAC | CAACAACAAC | AAACAACAAA | CAACATTACA | 1250 |
| | ATTACTATTT | ACAATTACAG | TCGACGGATC | AAGTGCAAAG | GTCCGCCTTG | 1300 |
| 35 | TTTCTCCTCT | GTCTCTTGAT | CTGACTAATC | TTGGTTTATG | ATTCGTTGAG | 1350 |

TAATTTTGGG GAAAGCTTCG TCCACAGTTT TTTTTCGAT GAACAGTGCC 1400
 GCAGTGGCGC TGATCTTGTA TGCTATCCTG CAATCGTGGT GAACTTATGT 1450
 CTTTTATATC CTTCACTACC ATGAAAAGAC TAGTAATCTT TCTCGATGTA 1500
 ACATCGTCCA GCACTGCTAT TACCGTGTGG TCCATCCGAC AGTCGGCTG 1550
 5 AACACATCAT ACGATATTGA GCAAAGATCG ATCTATCTTC CCTGTTCTTT 1600
 AATGAAAGAC GTCATTTTCA TCAGTATGAT CTAAGAATGT TGCAACTTGC 1650
 AAGGAGGCCGT TTCTTTCTTT GAATTTAACT AACTCGTTGA GTGGCCCTGT 1700
 TTCTCGGACG TAAGGCCCTT GCTGCTCCAC ACATGTCCAT TCGAATTTTA 1750
 CCGTGTTTAG CAAGGGCGAA AAGTTTGCAT CTTGATGATT TAGCTTGACT 1800
 10 ATGCGATTGC TTTCTGGAC CCGTGCAGCT GCGGACGGAT CCACCATGGC 1850
 ACTGCAGGTG GCATCTCTTA CAGACATAGG AGAATTGAGC AGTGTGGTTG 1900
 CTACTGGTTC TTGGTCTACT ACCTCGGCTA CTAATTTGAT GGAATTAAAC 1950
 ATTCATCCCA CCTCCTGTGC TATTCAGAAC GGATTGATAA CACAGACACC 2000
 ATTGAGTGTT TTAGCTCATG CTTTTCGAAG GTGGAGAGGA TCGTTGAAAA 2050
 15 TTTCCATCAT TTTCCGAGCG AGTTTGTTTA CCCGAGGACG AATCTTAGCC 2100
 GCTGCTGTGC CCGTTGCTAA GCGCAAAGGT ACCATGAGCC TTGACGAGAT 2150
 TAGTGGGTAT CATAATGTTT GCTGCTTATT GAATGGTCAG CAAACTACAT 2200
 TTGAAATTGGA AATCCCATAT TATTCTGTGG GCCAAGATTC TTTCTGTGAT 2250
 CGTGATGCTC TTTTGTATAT CTCTGCGCAC GATGGGAATT TTATGATTAC 2300
 20 TCGCTTGCAT CTCGTGATAC TGGATAAATT GGTAATGAGC GCTAATGCGA 2350
 GCAACAGCAT AAATTTTCC GTGACTCTTG GACCAGGTTT TGATTTGGAA 2400
 TTGAAATATC TTGCAGGAGT ACATGGGCAG CGCATAGTCC GCGAGTTGAA 2450
 GATGCAGTGA TCAACCTAGA CTTGTCCATC TTCTGGATTG GCCAACTTAA 2500
 TTAATGTATG AAATAAAAGG ATGCACACAT AGTGACATGC TAATCACTAT 2550
 25 AATGTGGGCA TCAAAGTTGT GTGTATGTG TAATTACTAG TTATCTGAAT 2600
 AAAAGAGAAA GAGATCATCC ATATTTCTTA TCCTAAATGA ATGTCACGTG 2650
 TCTTTATAAT TCTTTGATGA ACCAGATGCA TTTTATTAA CAAATCCATA 2700
 TACATATAAA TATTAATCAT ATATAATTAA TATCAATTGG GTTAGCAAAA 2750
 CAAATCTAGT CTAGGTGTGT TTTGCGAATT GCGGCCGCGA TCTGGGGAAT 2800
 30 TCGTAATCAT GGTCATAGCT GTTTCCTGTG TGAAATTGTT ATCCGCTCAC 2850
 AATTCCACAC AACATACGAG CCGGAAGCAT AAAGTGTAAG GCCTGGGGTG 2900
 CCTAATGAGT GAGCTAACTC ACATTAATTG CGTTGCGCTC ACTGCCCGCT 2950
 TTCCAGTCGG GAAACCTGTC GTGCCAGCTG CATTAAATGAA TCGGCCAACG 3000
 CGCGGGGAGA GCGGTTTTCG GTATTGGGCG CTCTCCGCT TCCTCGCTCA 3050
 35 CTGACTCGCT GCGCTCGGTC GTTCGGCTGC GCGGAGCGGT ATCAGCTCAC 3100

TCAAAGGCGG TAATACGGTT ATCCACAGAA TCAGGGGATA ACGCAGGAAA 3150
 GAACATGTGA GCAAAGGCC AGCAAAGGC CAGGAACCGT AAAAAGGCCG 3200
 CGTTGCTGGC GTTTTCCAT AGGCTCCGCC CCCCTGACGA GCATCACAAA 3250
 AATCGACGCT CAAGTCAGAG GTGGCGAAAC CCGACAGGAC TATAAAGATA 3300
 5 CCAGGCGTTT CCCCTGGAA GCTCCCTCGT GCGCTCTCCT GTTCCGACCC 3350
 TGCCGCTTAC CGGATACCTG TCCGCCTTTC TCCCTTCGGG AAGCGTGGCG 3400
 CTTTCTCATA GCTCACGCTG TAGGTATCTC AGTTCGGTGT AGGTCGTTCTG 3450
 CTCCAAGCTG GGCTGTGTGC ACGAACCCCC CGTTCAGCCC GACCGCTGCG 3500
 CCTTATCCGG TAACTATCGT CTTGAGTCCA ACCCGGTAAG ACACGACTTA 3550
 10 TCGCCACTGG CAGCAGCCAC TGGTAACAGG ATTAGCAGAG CGAGGTATGT 3600
 AGGCGGTGCT ACAGAGTTCT TGAAGTGGTG GCCTAACTAC GGCTACACTA 3650
 GAAGGACAGT ATTGGTATC TGCGCTCTGC TGAAGCCAGT TACCTTCGGA 3700
 AAAAGAGTTG GTAGCTCTTG ATCCGGCAAA CAAACCACCG CTGGTAGCGG 3750
 TGGTTTTTTT GTTGTCAAGC AGCAGATTAC GCGCAGAAAA AAAGGATCTC 3800
 15 AAGAAGATCC TTTGATCTTT TCTACGGGGT CTGACGCTCA GTGGAACGAA 3850
 AACTCAGCTT AAGGGATTTT GGTCAAGAGA TTATCAAAAA GGATCTTCAC 3900
 CTAGATCCTT TTAAATTAAT AATGAAGTTT TAAATCAATC TAAAGTATAT 3950
 ATGAGTAAAC TTGGTCTGAC AGTTACCAAT GCTTAATCAG TGAGGCACCT 4000
 ATCTCAGCGA TCTGTCTATT TCGTTTATCC ATAGTTGCGT GACTCCCCGT 4050
 20 CGTGTAGATA ACTACGATA GGGAGGGCTT ACCATCTGGC CCCAGTGCTG 4100
 CAATGATACC GCGAGACCCA CGCTCACCGG CTCCAGATTT ATCAGCAATA 4150
 AACCAGCCAG CCGGAAGGGC CGAGCGCAGA AGTGGTCCTG CAACTTTATC 4200
 CGCCTCCATC CAGTCTATTA ATTGTTGCCG GGAAGCTAGA GTAAGTAGTT 4250
 CGCCAGTTAA TAGTTTGCGC AACGTTGTTG CCATTGCTAC AGGCATCGTG 4300
 25 GTGTCACGCT CGTCGTTTGG TATGGCTTCA TTCAGTCCG GTTCCCAACG 4350
 ATCAAGGCGA GTTACATGAT CCCCCATGTT GTGCAAAAAA GCGGTTAGCT 4400
 CCTTCGGTCC TCCGATCGTT GTCAGAACTA AGTTGGCCGC AGTGTATCA 4450
 CTCATGGTTA TGGCAGCACT GCATAATTCT CTTACTGTCA TGCCATCCGT 4500
 AAGATGCTTT TCTGTGACTG GTGAGTACTC AACCAAGTCA TTCTGAGAAT 4550
 30 AGTGTATGCG GCGACCGAGT TGCTCTTGCC CGGCGTCAAT ACGGGATAAT 4600
 ACCGCGCCAC ATAGCAGAAC TTTAAAAGTG CTCATCATTG GAAACGTTT 4650
 TTCGGGGCGA AACTCTCAA GGATCTTACC GCTGTTGAGA TCCAGTTCGA 4700
 TGTAACCCAC TCGTGACCC AACTGATCTT CAGCATCTTT TACTTTCACC 4750
 AGCGTTTCTG GGTGAGCAAA AACAGGAAGG CAAAATGCCG CAAAAAGGG 4800
 35 AATAAGGGCG ACACGGAAT GTTGAATACT CATACTCTTC CTTTTTCAAT 4850

ATTATTGAAG CATTATCAG GGTATTGTC TCATGAGCGG ATACATATTT 4900
GAATGTATTT AGAAAAATAA ACAAATAGGG GTTCCGCGCA CATTTCCTCCG 4950
AAAAGTGCCA CCTGACGTCT AAGAAACCAT TATTATCATG ACATTAACCT 5000
ATAAAAATAG GCGTATCACG AGGCCCTTTC GTC 5033

5 (2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 45 bases

(B) TYPE: nucleotide

(C) STRANDEDNESS: single

10 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: viral RNA

(A) DESCRIPTION: RNA codons for first 15 amino acids at
5' end of MCDV coat protein 3 (CP3)

(iii) HYPOTHETICAL: No

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

CUG CAG GUG GCA UCU CUU ACA GAC AUA GGA GAA UUG AGC AGU GUG 45
Leu Gln Val Ala Ser Leu Thr Asp Ile Gly Asp Leu Ser Ser Val

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

20 (A) LENGTH: 20 bases

(B) TYPE: nucleotide

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: synthetic DNA

25 (A) DESCRIPTION: sequencing primer

(iii) HYPOTHETICAL: No

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GGTCTACTCA CGGCACGCCA

20

(2) INFORMATION FOR SEQ ID NO: 4:

30 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 11785 bases

(B) TYPE: nucleotide

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: viral RNA

(iii) HYPOTHETICAL: No

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

	XUGAAAAGGA	GGGUUAUAGAG	AUACCCUUCA	UAUAUUCUGC	GGAUGGCGUG	50
	CCGUGAGUAG	ACCUCGCGAC	GUUUCCAGAG	GGAAAUGGA	AAUGGUCCAU	100
5	GUAACACCAG	AUAUUUAUCU	GGUUGAGGAA	CAUGGUUUAG	UGGUAGAGAU	150
	AAACUCAACU	UUGUGUUGGA	CCCCGAUGCU	GUGAAAAGUA	AAUAAAGACA	200
	AGGCCACUUA	GCGAAGGAUA	UUCGAAGUAG	UGAUGAAAGG	AAGUGCAAUA	250
	AGUCAUGCCG	UAAGUCGCAA	UGCGCUAUAA	GUCAUGCCGU	AAGCCGCGUC	300
	GCCUGGAUUU	GCUAUUAGAA	UGUCCCUAGC	CGGUGAUAAAC	CUUGAGUCCC	350
10	CGUCAUAGGA	CUACUUUUGU	UUGCUUAGUA	AUACAUUGGG	ACCACCCGCA	400
	UGGAGCUCUG	AGCCUACCAU	ACAUAGUACA	UUUUCGAGG	GAUUGUCUUU	450
	UGAUA	AUG AUG CAG	ACA AAC AAC	CAA AAU CCC		485
		Met Met Gln	Thr Asn Asn	Asn Gln Asn	Pro	
	ACU CAA GGA	AGC AUU CCU	GAG AAC UCC	UCA CAA GAU	CGC AAC UUA	530
15	Thr Gln Gly	Ser Ile Pro	Glu Asn Ser	Ser Gln Asp	Arg Asn Leu	
	GGA GUG CCC	GCU GGA UAU	UCU UUA AGC	GUU GAG GAC	CCC UUC GGG	575
	Gly Val Pro	Ala Gly Tyr	Ser Leu Ser	Val Glu Asp	Pro Phe Gly	
	AAC CGG UCU	GAC UUU CAU	AUC CCA GUG	CAC CAA AUC	AUU CGG GAA	620
	Asn Arg Ser	Asp Phe His	Ile Pro Val	His Gln Ile	Ile Arg Glu	
20	GAG AUU GAU	CGU CCA AAU	UGG GUU CCU	AUA UGU UCA	AAC GAU UUU	665
	Glu Ile Asp	Arg Pro Asn	Trp Val Pro	Ile Cys Ser	Asn Asp Phe	
	CAU CUU AAC	AGU GAG GAU	UAU UGU GAG	GAG UGC GAA	UCU GAA CGG	710
	His Leu Asn	Ser Glu Asp	Tyr Cys Glu	Glu Cys Asp	Ser Asp Arg	
	AUC AAA AAU	UUC GAA AUA	UUC AGA UCA	CAG AAU UUG	AUU GAC CAA	755
25	Ile Lys Asn	Phe Asp Ile	Phe Arg Ser	Gln Asn Leu	Ile Asp Gln	
	CAC CUA AAU	CUC UGU ACU	GAU UCA AAG	GAU UGU GAU	CAU UUU UCU	800
	His Leu Asn	Leu Cys Thr	Asp Ser Lys	Asp Cys Asp	His Phe Ser	
	UGU UUU UCC	ACG AGU ACA	AGU UGC AGA	UUU UGC CCU	UUU UGC UUA	845
	Cys Phe Ser	Thr Ser Thr	Ser Cys Arg	Phe Cys Pro	Phe Cys Leu	
30	UUC AUU UUU	AAU UUG GAU	AAA UUU UAC	AAA CAA AAU	CUA UAU UUG	890
	Phe Ile Phe	Asn Leu Asp	Lys Phe Tyr	Lys Gln Asn	Leu Tyr Leu	
	AUU AGU CGU	CAG GCU CUA	GCU AGA UUG	UUC CAC GGA	AGC GCC GAA	935
	Ile Ser Arg	Gln Ala Leu	Ala Arg Leu	Phe His Gly	Ser Ala Asp	
	GAG UUA CUC	AGU AGA GCG	AUU UUC UUU	ACG UAU AAU	AUU UGU AUU	980
35	Glu Leu Leu	Ser Arg Ala	Ile Phe Phe	Thr Tyr Asn	Ile Cys Ile	

GAU GCA GAG GUG GUU GCU AAU AAU AGG AUU GGC UGU GAA UAU GUU 1025
 Asp Ala Glu Val Val Ala Asn Asn Arg Ile Gly Cys Asp Tyr Val
 AAG UUG UUU CAU CCA GAC CUU AGG CCU AGU AUU ACG UCU CCC CCU 1070
 Lys Leu Phe His Pro Asp Leu Arg Pro Ser Ile Thr Ser Pro Pro
 5 UAU GCU AGU GAU UGG GUU AUG UGU GAU AAU GCU AAA CAU CUU UUU 1115
 Tyr Ala Ser Asp Trp Val Met Cys Asp Asn Ala Lys His Leu Phe
 GAG UGU CUU GGC CUU GGU GAC ACG ACC AGA GGA CAC CUA UAU GGA 1160
 Glu Cys Leu Gly Leu Gly Asp Thr Thr Arg Gly His Leu Tyr Gly
 CUU AUU AGC GAG AAU GCA UAU UGG AAC GCC ACG UGC UCA AAA UGC 1205
 10 Leu Ile Ser Glu Asn Ala Tyr Trp Asn Ala Thr Cys Ser Lys Cys
 GGA GCC UGU UGU CAG GGA GCA AAU GCC CGU ACG GCG AUA CCG AUA 1250
 Gly Ala Cys Cys Gln Gly Ala Asn Ala Arg Thr Ala Ile Pro Ile
 GUG AUG GCG UUG CAG UAC UGC AGG GUG GAU GUG UAU UAU AGU GAG 1295
 Val Met Ala Leu Gln Tyr Cys Arg Val Asp Val Tyr Tyr Ser Glu
 15 UAC UAU UUA UAC CAC AUC UAC GCU CCG GAA GAG AGA AUG AAG AUU 1340
 Tyr Tyr Leu Tyr His Ile Tyr Ala Pro Asp Glu Arg Met Lys Ile
 GAU CAA CAG ACA GCA CAC UUG CUA CAC AGU AUA AUC CGA GGA GCA 1385
 Asp Gln Gln Thr Ala His Leu Leu His Ser Ile Ile Arg Gly Ala
 CCA GCA GUG GAU UGC UCU GAG UUA UCU CAG GAG CCA AUU CAC AGG 1430
 20 Pro Ala Val Asp Cys Ser Glu Leu Ser Gln Glu Pro Ile His Arg
 AUG GUA AUG GAU AGC UCA AAG UUA GUG GCA CUG GAU UCG ACA AUC 1475
 Met Val Met Asp Ser Ser Lys Leu Val Ala Leu Asp Ser Thr Ile
 AGG CAU CCU AAG AGC CAA GGA AGU UUG CUC GAU UCA GAA UGC GAU 1520
 Arg His Pro Lys Ser Gln Gly Ser Leu Leu Asp Ser Asp Cys Asp
 25 CAU GAG UUU AUU CUA AGA ACG UCC CAU GGU AUC AAA AUA CCG AUG 1565
 His Glu Phe Ile Leu Arg Thr Ser His Gly Ile Lys Ile Pro Met
 AGU AAG UCU UUA UUU AUA UCA UUU CUU ACC AUG GGA GCU UAU CAU 1610
 Ser Lys Ser Leu Phe Ile Ser Phe Leu Thr Met Gly Ala Tyr His
 GGG UAU GCU CAU GAU GAU CAG CAG GAG CAA AAU GCG AUA AUA UCU 1655
 30 Gly Tyr Ala His Asp Asp Gln Gln Glu Gln Asn Ala Ile Ile Ser
 UUU GGU GGG AUG CCC GGA GUC AAU UUG GCU UGU AAC AAA AAU UUC 1700
 Phe Gly Gly Met Pro Gly Val Asn Leu Ala Cys Asn Lys Asn Phe
 CUG AGA AUG CAU AAG UUG UUU UAU UCU GGA AGU UUU AGG CGC-AGA 1745
 Leu Arg Met His Lys Leu Phe Tyr Ser Gly Ser Phe Arg Arg Arg
 35 CCC CUG UUU AUG AGC CAA AUU CCC UCU ACG AAU GCC ACC GCU CAG 1790

Pro Leu Phe Met Ser Gln Ile Pro Ser Thr Asn Ala Thr Ala Gln
 UCC GGU UUU AAU GAU GAA GAA UUC GAA AGA UUG AUG GCU GAA GAG 1835
 Ser Gly Phe Asn Asp Asp Asp Phe Asp Arg Leu Met Ala Asp Glu
 GGU GUG CAU GUC AAA GUC GAG CGU CCA AUA GCA GAG AGG UUU GAU 1880
 5 Gly Val His Val Lys Val Glu Arg Pro Ile Ala Glu Arg Phe Asp
 UAU GAG GAC GUU AUU GAU AUU UAC GAU GAG ACC GAC CAC GAC AGG 1925
 Tyr Glu Asp Val Ile Asp Ile Tyr Asp Glu Thr Asp His Asp Arg
 ACA CGA GCU CUA GGC CUU GGC CAA GUA UUC GGA GGU UUG CUC AAA 1970
 Thr Arg Ala Leu Gly Leu Gly Gln Val Phe Gly Gly Leu Leu Lys
 10 GGA AUU UCU CAU UGU GUA GAU AGC CUA CAU AAG GUA UUU GAU UUC 2015
 Gly Ile Ser His Cys Val Asp Ser Leu His Lys Val Phe Asp Phe
 CCU CUG GAC CUG GCC AUA GAA GCA GCU CAG AAA ACU GGU GAU UGG 2060
 Pro Leu Asp Leu Ala Ile Asp Ala Ala Gln Lys Thr Gly Asp Trp
 CUU GAA GGA AAU AAA GCU GCA GUA GAU GAA ACU AAA AUU UGU GUG 2105
 15 Leu Asp Gly Asn Lys Ala Ala Val Asp Asp Thr Lys Ile Cys Val
 GGC UGU CCC GAG AUU CAA AAA GAU AUG AUC AGU UUC CAG AAU GAA 2150
 Gly Cys Pro Glu Ile Gln Lys Asp Met Ile Ser Phe Gln Asn Asp
 ACA AAA GAA GCU UUU GAA UUA AUA CGA UCA AGU AUA AAG AAG CUU 2195
 Thr Lys Asp Ala Phe Asp Leu Ile Arg Ser Ser Ile Lys Lys Leu
 20 UCC GAG GGC AUU GAC AAA AUC ACG AAG AUG AAU GCU ACG AAC UUU 2240
 Ser Glu Gly Ile Asp Lys Ile Thr Lys Met Asn Ala Thr Asn Phe
 GAA CGA AUC CUA GAC GGG AUU AAA CCA AUC GAG AGC AGG UUG ACA 2285
 Asp Arg Ile Leu Asp Gly Ile Lys Pro Ile Glu Ser Arg Leu Thr
 GAA CUU GAG AAC AAG GCA CCC GCU UCA GAC AGC AAA GCC AUG GAA 2330
 25 Asp Leu Glu Asn Lys Ala Pro Ala Ser Asp Ser Lys Ala Met Asp
 GCU CUG GUC CAG GCC GUG AAA GAC UUG AAA AUC AUG AAA GAG GCG 2375
 Ala Leu Val Gln Ala Val Lys Asp Leu Lys Ile Met Lys Glu Ala
 AUG CUC GAU CUA AAU CGA AGA CUG AGC AAG CUG GAA GGA AAG AAA 2420
 Met Leu Asp Leu Asn Arg Arg Leu Ser Lys Leu Asp Gly Lys Lys
 30 AGU GAU GGC CAG ACU ACU GAA GGG ACA GCG GGA GAG CAA CAA CCG 2465
 Ser Asp Gly Gln Thr Thr Asp Gly Thr Ala Gly Glu Gln Gln Pro
 AUC CCU AAG ACU CCA ACU CGA GUG AAG GCA AGA CCA GUU GUG AAG 2510
 Ile Pro Lys Thr Pro Thr Arg Val Lys Ala Arg Pro Val Val Lys
 CAA UCA GGA ACG AUA AUG GUA AAC GAA GAG AGC ACA GAA ACU UUC 2555
 35 Gln Ser Gly Thr Ile Met Val Asn Asp Glu Ser Thr Asp Thr Phe

AGG GAU AAU GAG AGU CGA GUG ACU GAC CCU AAC AGG AGC GAU AUG 2600
 Arg Asp Asn Glu Ser Arg Val Thr Asp Pro Asn Arg Ser Asp Met
 UUU GCU GCU GUU ACU GCA GAA UAC UUA GUU AAA UCG UUU ACA UGG 2645
 Phe Ala Ala Val Thr Ala Asp Tyr Leu Val Lys Ser Phe Thr Trp
 5 AAA GUU UCU GAU GGA CAA GAU AAA GUU UUG GCU GAC CUU GAU UUA 2690
 Lys Val Ser Asp Gly Gln Asp Lys Val Leu Ala Asp Leu Asp Leu
 CCU CAA GAC UUA UGG AAA UCC AAU UCC CGA UUG AGU GAU AUC AUG 2735
 Pro Gln Asp Leu Trp Lys Ser Asn Ser Arg Leu Ser Asp Ile Met
 GGG UAU UUC CAA UAU UAU GAU GCA ACC GGA AUC ACU UUU CGC AUA 2780
 10 Gly Tyr Phe Gln Tyr Tyr Asp Ala Thr Gly Ile Thr Phe Arg Ile
 ACG ACA ACA UGU GUU CCU AUG CAC GGU GGU ACU UUA UGU GCU GCU 2825
 Thr Thr Thr Cys Val Pro Met His Gly Gly Thr Leu Cys Ala Ala
 UGG GAU GCU AAU GGU UGC GCU ACA CGA CAA GGU AUA GCC ACA ACG 2870
 Trp Asp Ala Asn Gly Cys Ala Thr Arg Gln Gly Ile Ala Thr Thr
 15 GUU CAG CUG ACU GGU UUG CCC AAA ACA UUU AUU GAA GCU CAC AGC 2915
 Val Gln Leu Thr Gly Leu Pro Lys Thr Phe Ile Asp Ala His Ser
 UCA UCA GAA ACG AUA AUC GUG GUA AAG AAU UCC AAU AUA CAA UCC 2960
 Ser Ser Asp Thr Ile Ile Val Val Lys Asn Ser Asn Ile Gln Ser
 GCG AUU UGU CUA AGU GGA AGU GAG CAC UCG UUU GGG AGA AUG GGA 3005
 20 Ala Ile Cys Leu Ser Gly Ser Glu His Ser Phe Gly Arg Met Gly
 AUC CUG AAG AUC UGU UGC UUG AAU ACG UUG AAU GCG CCA AAG GAA 3050
 Ile Leu Lys Ile Cys Cys Leu Asn Thr Leu Asn Ala Pro Lys Asp
 GCU ACA CAG CAA GUG GCU GUG AAC GUC UGG AUU AAG UUU GAC GGA 3095
 Ala Thr Gln Gln Val Ala Val Asn Val Trp Ile Lys Phe Asp Gly
 25 GUU AAA UUU CAC GUU UAU UCU UUA AGG AAA AAU CCA GUC GUU UCG 3140
 Val Lys Phe His Val Tyr Ser Leu Arg Lys Asn Pro Val Val Ser
 CAA CUG CAG GUG GCA UCU CUU ACA GAC AUA GGA GAA UUG AGC AGU 3185
 Gln Leu Gln Val Ala Ser Leu Thr Asp Ile Gly Asp Leu Ser Ser
 GUG GUU GCU ACU GGU UCU UGG UCU ACU ACC UCG GCU ACU AAU UUG 3230
 30 Val Val Ala Thr Gly Ser Trp Ser Thr Thr Ser Ala Thr Asn Leu
 AUG GAA UUA AAC AUU CAU CCC ACC UCC UGU GCU AUU CAG AAC GGA 3275
 Met Asp Leu Asn Ile His Pro Thr Ser Cys Ala Ile Gln Asn Gly
 UUG AUA ACA CAG ACA CCA UUG AGU GUU UUA GCU CAU GCU UUU GCA 3320
 Leu Ile Thr Gln Thr Pro Leu Ser Val Leu Ala His Ala Phe Ala
 35 AGG UGG AGA GGA UCG UUG AAA AUU UCC AUC AUU UUC GGA GCG AGU 3365

Arg Trp Arg Gly Ser Leu Lys Ile Ser Ile Ile Phe Gly Ala Ser
 UUG UUU ACC CGA GGA CGA AUC UUA GCC GCU GCU GUG CCC GUU GCU 3410
 Leu Phe Thr Arg Gly Arg Ile Leu Ala Ala Ala Val Pro Val Ala
 AAG CGC AAA GGU ACC AUG AGC CUU GAC GAG AUU AGU GGG UAU CAU 3455
 5 Lys Arg Lys Gly Thr Met Ser Leu Asp Glu Ile Ser Gly Tyr His
 AAU GUU UGC UGC UUA UUG AAU GGU CAG CAA ACU ACA UUU GAA UUG 3500
 Asn Val Cys Cys Leu Leu Asn Gly Gln Gln Thr Thr Phe Asp Leu
 GAA AUC CCA UAU UAU UCU GUG GGC CAA GAU UCU UUC GUG UAC CGU 3545
 Asp Ile Pro Tyr Tyr Ser Val Gly Gln Asp Ser Phe Val Tyr Arg
 10 GAU GCU CUU UUU GAU AUC UCU GCG CAC GAU GGG AAU UUU AUG AUU 3590
 Asp Ala Leu Phe Asp Ile Ser Ala His Asp Gly Asn Phe Met Ile
 ACU CGC UUG CAU CUC GUG AUA CUG GAU AAA UUG GUA AUG AGC GCU 3635
 Thr Arg Leu His Leu Val Ile Leu Asp Lys Leu Val Met Ser Ala
 AAU GCG AGC AAC AGC AUA AAU UUU UCC GUG ACU CUU GGA CCA GGU 3680
 15 Asn Ala Ser Asn Ser Ile Asn Phe Ser Val Thr Leu Gly Pro Gly
 UCU GAU UUG GAA UUG AAA UAU CUU GCA GGA GUA CAU GGG CAG CGC 3725
 Ser Asp Leu Asp Leu Lys Tyr Leu Ala Gly Val His Gly Gln Arg
 AUA GUC CGC GAG UUG AAG AUG CAG GUU UCA UUG GGU CGG UCA UUU 3770
 Ile Val Arg Glu Leu Lys Met Gln Val Ser Leu Gly Arg Ser Phe
 20 GAG AAU GGA GUG CUU AUU GGU AGU GGC UUC GAC GAC UUG CUA CAA 3815
 Glu Asn Gly Val Leu Ile Gly Ser Gly Phe Asp Asp Leu Leu Gln
 AGA UGG AGU CAU UUG GUG UCC AUG CCU UUU AAU GCA AAA GGA GAC 3860
 Arg Trp Ser His Leu Val Ser Met Pro Phe Asn Ala Lys Gly Asp
 AGC GAU GAG AUC CAA GUC UUU GGC UAU AUC AUG ACU GUU GCC CCG 3905
 25 Ser Asp Glu Ile Gln Val Phe Gly Tyr Ile Met Thr Val Ala Pro
 GCG UAU CGU UCC CUU CCA GUC CAC UGC ACG CUG CUA AGU UGG UUU 3950
 Ala Tyr Arg Ser Leu Pro Val His Cys Thr Leu Leu Ser Trp Phe
 UCA CAA UUA UUC GUG CAG UGG AAA GGU GGU AUA AAG UAU AGA CUA 3995
 Ser Gln Leu Phe Val Gln Trp Lys Gly Gly Ile Lys Tyr Arg Leu
 30 CAC AUU GAU UCA GAA GAG CGC AGA UGG GGU GGA UUC AUC AAA GUU 4040
 His Ile Asp Ser Asp Glu Arg Arg Trp Gly Gly Phe Ile Lys Val
 UGG CAU GAC CCA AAU GGC UCU UUG GAU GAA GGG AAA GAA UUU GCU 4085
 Trp His Asp Pro Asn Gly Ser Leu Asp Asp Gly Lys Asp Phe Ala
 AAA GCG GAU AUU CUA UCG CCA CCA GCC GGA GCU AUG GUU CGU UAU 4130
 35 Lys Ala Asp Ile Leu Ser Pro Pro Ala Gly Ala Met Val Arg Tyr

UGG AAC UAU UUA AAU GGA GAC UUG GAG UUU ACA GUA CCA UUU UGU 4175
 Trp Asn Tyr Leu Asn Gly Asp Leu Glu Phe Thr Val Pro Phe Cys
 GCU AGA ACC AGU ACG CUG UUC AUA CCA AAA GCU AUG AUU GCC ACC 4220
 Ala Arg Thr Ser Thr Leu Phe Ile Pro Lys Ala Met Ile Ala Thr
 5 GAU UCA AAG UCA UGG AUU CUG AAC UAC AAC GGU ACA UUG AAU UUC 4265
 Asp Ser Lys Ser Trp Ile Leu Asn Tyr Asn Gly Thr Leu Asn Phe
 GCG UAC CAA GGA GUA GAU GAC UUC ACA AUU ACA GUG GAA ACA AGU 4310
 Ala Tyr Gln Gly Val Asp Asp Phe Thr Ile Thr Val Asp Thr Ser
 GCA GCC GAC GAC UUU GAA UUU CAC GUU CGA ACA GUU GCA CCC CGC 4355
 10 Ala Ala Asp Asp Phe Asp Phe His Val Arg Thr Val Ala Pro Arg
 GCU GGA AAG GUC AAC GAA GCU UUU GCC AAA UUG GAG UAC GCU UCU 4400
 Ala Gly Lys Val Asn Asp Ala Phe Ala Lys Leu Glu Tyr Ala Ser
 GAU UUA AAG GAU AUC AAA GAA UCU CUG ACA UCU UCC ACU CGU UUG 4445
 Asp Leu Lys Asp Ile Lys Asp Ser Leu Thr Ser Ser Thr Arg Leu
 15 AAA GGG CCU CAU UAU AAA ACG AAA AUU ACC UCA AUA GAG CCA AAU 4490
 Lys Gly Pro His Tyr Lys Thr Lys Ile Thr Ser Ile Glu Pro Asn
 AAA AUU GAU GAA AAU GAG UCC UCA CGU GGU AAA GAU AAC AAG UCA 4535
 Lys Ile Asp Asp Asn Glu Ser Ser Arg Gly Lys Asp Asn Lys Ser
 AAU UCG AAA UUU GAG GAC UUA CUC AAU GCA ACA GCU CAG AUG GAU 4580
 20 Asn Ser Lys Phe Glu Asp Leu Leu Asn Ala Thr Ala Gln Met Asp
 UUU GAU CGA GCC ACA GCG AAC GUU GGG UGU GUG CCA UUC UCC AUU 4625
 Phe Asp Arg Ala Thr Ala Asn Val Gly Cys Val Pro Phe Ser Ile
 GCA AAG ACA GCA AAG GUG CUU UCG GAA CGC GAG ACG UGU AAG AAG 4670
 Ala Lys Thr Ala Lys Val Leu Ser Asp Arg Glu Thr Cys Lys Lys
 25 AUG GCA GAU GUG UUA GAU UUC ACA CAC UCA UGU UUG AAC UUA GAC 4715
 Met Ala Asp Val Leu Asp Phe Thr His Ser Cys Leu Asn Leu Asp
 AGU CAA CCU GCG GCG GCA AGA UUA GCA GCG GCC AUU UCU CAA AUA 4760
 Ser Gln Pro Ala Ala Ala Arg Leu Ala Ala Ala Ile Ser Gln Ile
 GCA CCU AUU AUG GAG AGC AUC GGU AGA ACC ACU CAA AGC GUA GAG 4805
 30 Ala Pro Ile Met Glu Ser Ile Gly Arg Thr Thr Gln Ser Val Glu
 GAA AAA UUG GCU UCU GUG GAU ACA UUU AGG GAC AAA AUC AUG GCU 4850
 Asp Lys Leu Ala Ser Val Asp Thr Phe Arg Asp Lys Ile Met Ala
 CUA AUU UCA AAC GUG CUU GGG GAU ACU CUA CCU GGA CUG GCC AUU 4895
 Leu Ile Ser Asn Val Leu Gly Asp Thr Leu Pro Gly Leu Ala Ile
 35 GCU GAC UUC AAA AAA GGA AAA UAU GUG UGG GCC UCG UUC CUG ACA 4940

Ala Asp Phe Lys Lys Gly Lys Tyr Val Trp Ala Ser Phe Leu Thr
 AUG AUA GCC GCU UGC GUA GUA GCU UGG GCU GCC ACU AGC AAG AAA 4985
 Met Ile Ala Ala Cys Val Val Ala Trp Ala Ala Thr Ser Lys Lys
 AGC UUC UUG AAA AGA UUU GCA GUG GUA GCU AUG AUA AUU UGG AGC 5030
 5 Ser Phe Leu Lys Arg Phe Ala Val Val Ala Met Ile Ile Trp Ser
 CCA UUU CUC GCA AGU AAA AUA UGG GCG CUU GGU ACA UGG AUU AGG 5075
 Pro Phe Leu Ala Ser Lys Ile Trp Ala Leu Gly Thr Trp Ile Arg
 AAG AGC UGG AGU AAG CUU UGG CCU AAG UCA GAC UCA UGC CGA CAA 5120
 Lys Ser Trp Ser Lys Leu Trp Pro Lys Ser Asp Ser Cys Arg Gln
 10 CAC UCU UUG GCA GGC CUG UGU GAA AGU GUG UUC ACA UCA UUC AAG 5165
 His Ser Leu Ala Gly Leu Cys Asp Ser Val Phe Thr Ser Phe Lys
 GAU UUC CCU GAC UGG UUU AAA UCA GGA GGA AUC ACG AUU GUG ACG 5210
 Asp Phe Pro Asp Trp Phe Lys Ser Gly Gly Ile Thr Ile Val Thr
 CAA GUU UGC ACA GUA UUA CUG ACG AUA GUG AGU CUG AUU ACA CUU 5255
 15 Gln Val Cys Thr Val Leu Leu Thr Ile Val Ser Leu Ile Thr Leu
 GGA ACU AUA CCA AGC ACG AAA CAA AAU GCU ACG UUC GCA GAC AAA 5300
 Gly Thr Ile Pro Ser Thr Lys Gln Asn Ala Thr Phe Ala Asp Lys
 UUU AAA GAA UUU GGU AAC AUG AGC AGA GCU ACA ACG UCA AUA GCU 5345
 Phe Lys Asp Phe Gly Asn Met Ser Arg Ala Thr Thr Ser Ile Ala
 20 GCA GGU UAC AAG ACG AUA UCA GAG CUG UGU UCG AAA UUC ACC AAU 5390
 Ala Gly Tyr Lys Thr Ile Ser Glu Leu Cys Ser Lys Phe Thr Asn
 UAC UUG GCU GUA ACC UUC UUU GGG GCG CAA GUU GAU GAC GAU GCU 5435
 Tyr Leu Ala Val Thr Phe Phe Gly Ala Gln Val Asp Asp Asp Ala
 UUC AAG GGU UUG GUA GCG UUC AAC GUU AAG GAA UGG AUU CUU GAA 5480
 25 Phe Lys Gly Leu Val Ala Phe Asn Val Lys Asp Trp Ile Leu Asp
 GUG AAA AAC CUG UCU CUU GAG GAA AAC AAA UUU AGU GGU UUU GGU 5525
 Val Lys Asn Leu Ser Leu Glu Asp Asn Lys Phe Ser Gly Phe Gly
 GGU GAU GAG CAU CUU GUC AAG GUU AGA CAU UUA UAU GAU AAA UCU 5570
 Gly Asp Glu His Leu Val Lys Val Arg His Leu Tyr Asp Lys Ser
 30 GUG GAA AUA ACC UAU AAG UUG CUC CAG AAA AAU CGA GUU CCC AUU 5615
 Val Asp Ile Thr Tyr Lys Leu Leu Gln Lys Asn Arg Val Pro Ile
 GCU AUG CUU CCU AUC AUC CGA GAC ACG UGU AAG AAG UGC GAG GAU 5660
 Ala Met Leu Pro Ile Ile Arg Asp Thr Cys Lys Lys Cys Glu Asp
 UUG CUA AAC GAG AGU UAU ACU UAC AAA GGU AUG AAA ACU CCG CGC 5705
 35 Leu Leu Asn Glu Ser Tyr Thr Tyr Lys Gly Met Lys Thr Pro Arg

GUG GAC CCA UUC UAU AUA UGC CUU UUU GGA GCA CCU GGA GUU GGC 5750
 Val Asp Pro Phe Tyr Ile Cys Leu Phe Gly Ala Pro Gly Val Gly
 AAG UCC ACA GUG GCA UCG AUG AUU GUU GAC GAU UUG UUG GAU GCU 5795
 Lys Ser Thr Val Ala Ser Met Ile Val Asp Asp Leu Leu Asp Ala
 5 AUG GGC GAA CCU AAG GUU GAU AGG AUC UAU ACG CGA UGC UGU UCU 5840
 Met Gly Asp Pro Lys Val Asp Arg Ile Tyr Thr Arg Cys Cys Ser
 GAU CAA UAU UGG AGC AAU UAU CAC CAC GAG CCA GUU AUU UGU UAU 5885
 Asp Gln Tyr Trp Ser Asn Tyr His His Glu Pro Val Ile Cys Tyr
 GAC GAC UUG GGG GCA AUC AGC AGA CCA GCG AGU UUA UCA GAC UAU 5930
 10 Asp Asp Leu Gly Ala Ile Ser Arg Pro Ala Ser Leu Ser Asp Tyr
 GGG GAG AUA AUG GGA AUC AAA UCG AAC AGA CCA UAC UCC CUA CCU 5975
 Gly Glu Ile Met Gly Ile Lys Ser Asn Arg Pro Tyr Ser Leu Pro
 AUG GCU GCU GUU GAU GAG AAA GGA AGG CAU UGU UUA UCG CGA UAC 6020
 Met Ala Ala Val Asp Glu Lys Gly Arg His Cys Leu Ser Arg Tyr
 15 CUC AUU GCU UGU ACA AAU UUA ACC CAU CUG GAC GAU ACG GGC GAU 6065
 Leu Ile Ala Cys Thr Asn Leu Thr His Leu Asp Asp Thr Gly Asp
 GUG AAA ACA AAG GAU GCC UAC UAU CGC AGA AUC AAU GUC CCA GUG 6110
 Val Lys Thr Lys Asp Ala Tyr Tyr Arg Arg Ile Asn Val Pro Val
 ACA GUG ACG AGA GAA GUA ACC GCC AUG AUG AAC CCC GAG GAC CCA 6155
 20 Thr Val Thr Arg Asp Val Thr Ala Met Met Asn Pro Glu Asp Pro
 ACU GAU GGA CUA CGU UUC ACC GUG GAG CAA GUG CUU GAU GGA GGU 6200
 Thr Asp Gly Leu Arg Phe Thr Val Glu Gln Val Leu Asp Gly Gly
 AGA UGG AUU AAU GUU ACU GAA AGC CGU CUC CUC AAU GGA AGG AUG 6245
 Arg Trp Ile Asn Val Thr Asp Ser Arg Leu Leu Asn Gly Arg Met
 25 CCA UUC AGG GCU GAA GAU CUC AUG AAC AUG AAC UAC AGU UAC UUU 6290
 Pro Phe Arg Ala Asp Asp Leu Met Asn Met Asn Tyr Ser Tyr Phe
 AUG GAG UUU CUC AAG AUG UAU GCU GCU UUA UAU AUG GAA AAU CAA 6335
 Met Glu Phe Leu Lys Met Tyr Ala Ala Leu Tyr Met Asp Asn Gln
 AAC AUG UUG GUG GCA AAA UUG AGA GGA ACA GAG AUC CCA GAA UCA 6380
 30 Asn Met Leu Val Ala Lys Leu Arg Gly Thr Glu Ile Pro Asp Ser
 CGU AGU UCA GAG AAU GAA GAA CUU GAA UUC GAU UAU UUG GCU ACA 6425
 Arg Ser Ser Glu Asn Asp Asp Leu Asp Phe Asp Tyr Leu Ala Thr
 GCU CAG AUG GAC CAU ACA GUG ACA UUU GGG GAA CUA GUU ACC AAA 6470
 Ala Gln Met Asp His Thr Val Thr Phe Gly Asp Leu Val Thr Lys
 35 UUC AAC UCG UAU AAG CUU ACU GGG AAA CAA UGG AAC AAG AGG CUC 6515

Phe Asn Ser Tyr Lys Leu Thr Gly Lys Gln Trp Asn Lys Arg Leu
 UGU GAA CUU GGA UGG ACA UCU CUA GAC GGA UGG AAC ACG AAC AAG 6560
 Cys Asp Leu Gly Trp Thr Ser Leu Asp Gly Trp Asn Thr Asn Lys
 AUU AUG AGA UUC GAC GAU CUA GUU GCC GGA UUC UGU GGU UGC UCA 6605
 5 Ile Met Arg Phe Asp Asp Leu Val Ala Gly Phe Cys Gly Cys Ser
 AGG AAU GAG AAU UGC AAU UUU GAC UUC UAU CAU CAG AGA CUU CAA 6650
 Arg Asn Glu Asn Cys Asn Phe Asp Phe Tyr His Gln Arg Leu Gln
 GCA UGU UUG AAC AAG AAA GGG UUU GCU CCC GCA UAU CAA UAU UUC 6695
 Ala Cys Leu Asn Lys Lys Gly Phe Ala Pro Ala Tyr Gln Tyr Phe
 10 AAC CUU CAC AAG UUG AAU UCA GAC ACC CAG AAG ACA GAG CUC AAG 6740
 Asn Leu His Lys Leu Asn Ser Asp Thr Gln Lys Thr Glu Leu Lys
 CUU AAA UGC GGG ACA ACU GCU GAA GAU UUA UUC AGA CAA GCU GAC 6785
 Leu Lys Cys Gly Thr Thr Ala Asp Asp Leu Phe Arg Gln Ala Asp
 UUG AUG GUC AUA UUC UCC UAC CUC UUA UUU GUU GCG AGA AUU GGG 6830
 15 Leu Met Val Ile Phe Ser Tyr Leu Leu Phe Val Ala Arg Ile Gly
 GUG AGU GGA UCU CAU GUG UGU CUG UCA UAU AAC AUG UUG AAC GUC 6875
 Val Ser Gly Ser His Val Cys Leu Ser Tyr Asn Met Leu Asn Val
 AAG GAU GUC AAG GAU UUU GAG AUA UGC AGG GAG AAC GUU CUU GAU 6920
 Lys Asp Val Lys Asp Phe Glu Ile Cys Arg Glu Asn Val Leu Asp
 20 UUG UCC AGA AAA ACU ACA AUC GAC GGU GAA GAA UGC UAU AUC UGG 6965
 Leu Ser Arg Lys Thr Thr Ile Asp Gly Asp Asp Cys Tyr Ile Trp
 AAU UUU AUU UCU GAU AUC UUC CCA CGC AUU GUG GCU AAG UAC AAC 7010
 Asn Phe Ile Ser Asp Ile Phe Pro Arg Ile Val Ala Lys Tyr Asn
 UGU GUU GUG CUU AAC GAC GGA GAG AAG AGA UAC AUC UUC GUG ACU 7055
 25 Cys Val Val Leu Asn Asp Gly Glu Lys Arg Tyr Ile Phe Val Thr
 GAC AGC GCG CCC ACU AGG AUC UUU CCC GAU UUG GCU UGG UCA GAU 7100
 Asp Ser Ala Pro Thr Arg Ile Phe Pro Asp Leu Ala Trp Ser Asp
 CUU AUU UCC GGC AAG CAA GUU GUG AGU CCA AAC AUU AUC AAA GUG 7145
 Leu Ile Ser Gly Lys Lys Gln Val Val Ser Pro Asn Ile Ile Lys Val
 30 GCU GGA GAA ACC AAG UCG AAA ACC AUU GCC CCU CUG CUA GCA GAU 7190
 Ala Gly Asp Thr Lys Ser Lys Thr Ile Ala Pro Leu Leu Ala Asp
 UCC UAC AAG GUU UUC AAG GAU CCG AAG GCA UGG CUU GAG AGG AAC 7235
 Ser Tyr Lys Val Phe Lys Asp Pro Lys Ala Trp Leu Glu Arg-Asn
 AAA GAA UUG AAA GCA GCU CUA GAA ACA GAA GAA UAU AUC GCU CUC 7280
 35 Lys Asp Leu Lys Ala Ala Leu Asp Thr Asp Asp Tyr Ile Ala Leu

CUC UUU GCU GUU GCA UGU GAA GCU GGU AGA UUC ACU CAA AUU UUA 7325
 Leu Phe Ala Val Ala Cys Asp Ala Gly Arg Phe Thr Gln Ile Leu
 GAC AAA CCU CCC AGU AGA CGC AAG AUU UUA AAU AUG UCC GAA AGG 7370
 Asp Lys Pro Pro Ser Arg Arg Lys Ile Leu Asn Met Ser Asp Arg
 5 UAU AAU GCA UAU AUU GAA CAG GAA AAA GGG CUG AUU GGG AGA CUU 7415
 Tyr Asn Ala Tyr Ile Asp Gln Asp Lys Gly Leu Ile Gly Arg Leu
 UCU AAA CCA GCA AAG AUA UGC UUA GCC AUA GGA ACU GGA GUU GCG 7460
 Ser Lys Pro Ala Lys Ile Cys Leu Ala Ile Gly Thr Gly Val Ala
 AUC UUU GGG GCC CUA GCA GGC AUU GGA GUG GGU UUG UUU AAG CUG 7505
 10 Ile Phe Gly Ala Leu Ala Gly Ile Gly Val Gly Leu Phe Lys Leu
 AUA GCU CAC UUC AAC AAA GAU GAA GAA GAG GUA GAC GAA AUU GAA 7550
 Ile Ala His Phe Asn Lys Asp Asp Asp Glu Val Asp Asp Ile Asp
 UUU GAU AUA CUC UCC CCA GAG AUG AGC GGU UCG CAC GAA UCC GGC 7595
 Phe Asp Ile Leu Ser Pro Glu Met Ser Gly Ser His Asp Ser Gly
 15 CAA CAU ACC ACG AGG UAC GUC ACG AAG GAG CGA GUU CCA UCC AAA 7640
 Gln His Thr Thr Arg Tyr Val Thr Lys Glu Arg Val Pro Ser Lys
 CCA GCA AGG AGG CAA CAU GAA UUU GAU CUA AUG UUC GAU AAU CUA 7685
 Pro Ala Arg Arg Gln His Asp Phe Asp Leu Met Phe Asp Asn Leu
 CCC ACU CCA CAA GUU GAA GAG CUA AAG AGU GAG AUG ACC UGC GCC 7730
 20 Pro Thr Pro Gln Val Asp Glu Leu Lys Ser Glu Met Thr Cys Ala
 AGU GCC AGU GAU GAG CAU AAG ACU CAG UAU GUU AAA AGA AGA GUG 7775
 Ser Ala Ser Asp Glu His Lys Thr Gln Tyr Val Lys Arg Arg Val
 GGA CCU GUA AGC AAA CGU AAG GAU GCU UCG GUA GCA GAA AUU AGU 7820
 Gly Pro Val Ser Lys Arg Lys Asp Ala Ser Val Ala Asp Ile Ser
 25 GGA GCU CAU GCG AGU GAU CAG CAU CAU ACA GAA UAC UUG AAA GCA 7865
 Gly Ala His Ala Ser Asp Gln His His Thr Asp Tyr Leu Lys Ala
 CGC GUU CCA CUC AUG AAA AGA AUA GCU ACC AAA GAG AGC UAU GUU 7910
 Arg Val Pro Leu Met Lys Arg Ile Ala Thr Lys Glu Ser Tyr Val
 GUA ACU UAC GAU GAC GAA CCC AGC UCU CAU AUU UCC CUA GUU CGC 7955
 30 Val Thr Tyr Asp Asp Asp Pro Ser Ser His Ile Ser Leu Val Arg
 AGG AUC CGA CGU ACA CGA CUG GCA AGA GCC AUC AAG CAA AUG GCA 8000
 Arg Ile Arg Arg Thr Arg Leu Ala Arg Ala Ile Lys Gln Met Ala
 GUC CUG GAG GAC UUC CCA UCU ACC UUG GAA GAG AUA CGA CUU UGG 8045
 Val Leu Glu Asp Phe Pro Ser Thr Leu Asp Glu Ile Arg Leu Trp
 35 AGA CAA AAC GCU GCA AAU AAA GGG GUU AUU GUU CCG AAG UAC UCA 8090

Arg Gln Asn Ala Ala Asn Lys Gly Val Ile Val Pro Lys Tyr Ser
 ACA AGU GGG AAA UUC UUC AGU GGC UUG UUG GAU GAU GAA GAA GAA 8135
 Thr Ser Gly Lys Phe Phe Ser Gly Leu Leu Asp Asp Asp Asp Asp
 GAA CCU CAG AAU GUG AAU AUG UUG AAC GAA GAG GAC AUU GAG GUA 8180
 5 Asp Pro Gln Asn Val Asn Met Leu Asn Asp Glu Asp Ile Glu Val
 GAU AAG CGA AUG UUU GAG AAG AUU UCU GAG GUU AUA AGC GUG AUU 8225
 Asp Lys Arg Met Phe Glu Lys Ile Ser Glu Val Ile Ser Val Ile
 CAA CCC AGA AAG AAU GAG CUG GAA AGA AUG AUU GAG GAA GGC GUA 8270
 Gln Pro Arg Lys Asn Glu Leu Asp Arg Met Ile Glu Asp Gly Val
 10 CAC CAC AAG GUC GUA AAG CAG GCA AGG GUU AAC GAC AAG GGC UUA 8315
 His His Lys Val Val Lys Gln Ala Arg Val Asn Asp Lys Gly Leu
 GCC AAA GAC CCC AAC AUG GUG ACU AUC UUG ACG GAC AAA UUA AUU 8360
 Ala Lys Asp Pro Asn Met Val Thr Ile Leu Thr Asp Lys Leu Ile
 AAU AUU AGU GCG GUG AUC GUC AAU UUA ACG CCG ACA CGC CGG GCA 8405
 15 Asn Ile Ser Ala Val Ile Val Asn Leu Thr Pro Thr Arg Arg Ala
 UAC AUG AAC GUG GUA CGU CUU AUA GGC ACU AUA GUU GUU UGC CCA 8450
 Tyr Met Asn Val Val Arg Leu Ile Gly Thr Ile Val Val Cys Pro
 GCC CAC UAC UUG GAA GCU UUA GAG GAA GGA GAU GAG CUG UAU UUC 8495
 Ala His Tyr Leu Asp Ala Leu Glu Asp Gly Asp Glu Leu Tyr Phe
 20 AUU UGC UUC UCA UUG GUU AUC AAG CUC ACU UUU GAU CCA AGU AGA 8540
 Ile Cys Phe Ser Leu Val Ile Lys Leu Thr Phe Asp Pro Ser Arg
 GUG ACU CUC GUG AAU AGC CAG CAG GAU UUG AUG GUU UGG GAU CUU 8585
 Val Thr Leu Val Asn Ser Gln Gln Asp Leu Met Val Trp Asp Leu
 GGG AAC AUG GUA CCA CCC UCA AUU GAU ACU CUU AAA AUG AUA CCU 8630
 25 Gly Asn Met Val Pro Pro Ser Ile Asp Thr Leu Lys Met Ile Pro
 ACG CUU GAA GAC UGG GAU CAC UUU CAG GAU GGA CCA GGA GCC UUU 8675
 Thr Leu Asp Asp Trp Asp His Phe Gln Asp Gly Pro Gly Ala Phe
 GCU GUU ACG AAA UAU AAC UCG AAA UUC CCA ACC AAU UAU AUC AAC 8720
 Ala Val Thr Lys Tyr Asn Ser Lys Phe Pro Thr Asn Tyr Ile Asn
 30 ACA CUG ACU AUG AUU GAG AGG AUU AGG GCA AAU ACU CAG AAU CCC 8765
 Thr Leu Thr Met Ile Glu Arg Ile Arg Ala Asn Thr Gln Asn Pro
 ACG GGU UGU UAU UCC AUG AUG GGC UCC CAA CAU ACA AUC ACC ACA 8810
 Thr Gly Cys Tyr Ser Met Met Gly Ser Gln His Thr Ile Thr-Thr
 GGA UUG CGA UAU CAA AUG UUC UCU CUU GAU GGA UUC UGC GGU GGG 8855
 35 Gly Leu Arg Tyr Gln Met Phe Ser Leu Asp Gly Phe Cys Gly Gly

	UUA	AUC	CUG	AGA	GCC	AGC	ACA	AAC	AUG	GUG	AGA	AAG	GUC	GUC	GGG	8900
	Leu	Ile	Leu	Arg	Ala	Ser	Thr	Asn	Met	Val	Arg	Lys	Val	Val	Gly	
	AUC	CAC	GUU	GCU	GGA	AGC	CAG	AAU	CAC	GCU	AUG	GGA	UAU	GCA	GAG	8945
	Ile	His	Val	Ala	Gly	Ser	Gln	Asn	His	Ala	Met	Gly	Tyr	Ala	Glu	
5	UGC	CUU	AUU	GCA	GAA	GAU	UUA	CGG	GCU	GCA	GUG	GCG	AGA	UUG	GCG	8990
	Cys	Leu	Ile	Ala	Asp	Asp	Leu	Arg	Ala	Ala	Val	Ala	Arg	Leu	Ala	
	CUA	GAU	CCU	AGA	AGC	ACC	AUC	CAG	GCA	AGU	CUG	AAA	GGU	AGG	AUU	9035
	Leu	Asp	Pro	Arg	Ser	Thr	Ile	Gln	Ala	Ser	Leu	Lys	Gly	Arg	Ile	
	GAU	GCU	GUU	UCU	AAA	CAA	UGU	GGU	UUA	GAC	AGA	GCU	CUG	GGU	ACG	9080
10	Asp	Ala	Val	Ser	Lys	Gln	Cys	Gly	Leu	Asp	Arg	Ala	Leu	Gly	Thr	
	AUA	GGA	UGU	CAC	GGG	AAA	GUU	GCC	UCU	GAA	GAU	AUU	ACA	AGU	GCC	9125
	Ile	Gly	Cys	His	Gly	Lys	Val	Ala	Ser	Asp	Asp	Ile	Thr	Ser	Ala	
	GCC	ACG	AAA	ACU	UCC	AUA	AGA	AAG	UCA	AGA	AUA	CAU	GGU	CUA	GUG	9170
	Ala	Thr	Lys	Thr	Ser	Ile	Arg	Lys	Ser	Arg	Ile	His	Gly	Leu	Val	
15	GGU	GAG	AUU	AGA	ACU	GAG	CCU	UCA	AUU	UUA	CAC	GCU	CAU	GAU	CCC	9215
	Gly	Glu	Ile	Arg	Thr	Glu	Pro	Ser	Ile	Leu	His	Ala	His	Asp	Pro	
	CGA	CUG	CCU	AAA	GAC	AAG	AUU	GGG	AAA	UGG	GAC	CCG	GUU	AUU	GAG	9260
	Arg	Leu	Pro	Lys	Asp	Lys	Ile	Gly	Lys	Trp	Asp	Pro	Val	Ile	Glu	
	GCA	UCA	AUG	AAG	UAU	GGU	UCG	AGA	AUC	ACA	CCG	UUC	CCU	GUA	GAC	9305
20	Ala	Ser	Met	Lys	Tyr	Gly	Ser	Arg	Ile	Thr	Pro	Phe	Pro	Val	Asp	
	CAA	AUU	CUG	GAA	GUG	GAG	GAU	CAU	CUU	UCU	AAA	AUG	UUG	GCC	AAU	9350
	Gln	Ile	Leu	Asp	Val	Glu	Asp	His	Leu	Ser	Lys	Met	Leu	Ala	Asn	
	UGU	GAG	AAU	UCA	AAA	AAC	AAG	CGG	CAG	GUU	AAU	AAU	CUA	GAA	AUA	9395
	Cys	Glu	Asn	Ser	Lys	Asn	Lys	Arg	Gln	Val	Asn	Asn	Leu	Asp	Ile	
25	GGG	AUU	AAU	GGA	AUU	GAC	CAG	UCG	GAU	UAU	UGG	CAA	CAG	AUA	GAA	9440
	Gly	Ile	Asn	Gly	Ile	Asp	Gln	Ser	Asp	Tyr	Trp	Gln	Gln	Ile	Asp	
	AUG	GAU	ACU	UCA	AGU	GGU	UGG	CCA	UAC	GCU	AAG	CGU	AAA	CCU	GUU	9485
	Met	Asp	Thr	Ser	Ser	Gly	Trp	Pro	Tyr	Ala	Lys	Arg	Lys	Pro	Val	
	GGG	GCA	GCU	GGA	AAG	AAA	UGG	CUA	UUC	GAG	CAA	GAC	GGC	ACA	UAU	9530
30	Gly	Ala	Ala	Gly	Lys	Lys	Trp	Leu	Phe	Glu	Gln	Asp	Gly	Thr	Tyr	
	CCC	UCC	GGA	AAA	CCU	CGA	UAU	GUA	UUU	GGA	GAU	GCC	GGG	UUG	AUU	9575
	Pro	Ser	Gly	Lys	Pro	Arg	Tyr	Val	Phe	Gly	Asp	Ala	Gly	Leu	Ile	
	GAG	AGC	UAU	AAC	UCG	AUG	CUU	GGU	GAG	GCG	AAG	CAA	GGC	AUU-AGU		9620
	Glu	Ser	Tyr	Asn	Ser	Met	Leu	Gly	Glu	Ala	Lys	Gln	Gly	Ile	Ser	
35	CCC	ACU	GUC	GUC	ACA	AUU	GAG	UGC	GCA	AAA	GAU	GAG	AGG	CGG	AAG	9665

Pro Thr Val Val Thr Ile Glu Cys Ala Lys Asp Glu Arg Arg Lys
 CUU AAU AAG AUA UAU GAG AAA CCC GCC ACU CGG ACG UUC ACC AUA 9710
 Leu Asn Lys Ile Tyr Glu Lys Pro Ala Thr Arg Thr Phe Thr Ile
 CUG CCA CCU GAG AUU AAU AUU UUA UUC AGG CAG UAU UUC GGA GAU 9755
 5 Leu Pro Pro Glu Ile Asn Ile Leu Phe Arg Gln Tyr Phe Gly Asp
 UUU GCA GCG AUG GUA AUG ACA UGU AGA GCC AAG CUU UUC UGU CAA 9800
 Phe Ala Ala Met Val Met Thr Cys Arg Ala Lys Leu Phe Cys Gln
 GUU GGC AUC AAC CCA GAG UCA AUG GAG UGG GGU GAU CUC AUG CUA 9845
 Val Gly Ile Asn Pro Glu Ser Met Glu Trp Gly Asp Leu Met Leu
 10 GGU CUA AAG GAG AAA UCA ACU AAG GGA UUU GCA GGA GAU UAU UCG 9890
 Gly Leu Lys Glu Lys Ser Thr Lys Gly Phe Ala Gly Asp Tyr Ser
 AAG UUC GAU GGA AUC GGA GAC CCC CAG AUU UAU CAU UCA AUU ACC 9935
 Lys Phe Asp Gly Ile Gly Asp Pro Gln Ile Tyr His Ser Ile Thr
 CAA GUA GUC AAC AAC UGG UAU AAC GAU GGG GAA GAA AAU GCG ACU 9980
 15 Gln Val Val Asn Asn Trp Tyr Asn Asp Gly Asp Asp Asn Ala Thr
 AUC AGG CAU GCU CUG AUA AGU AGC AUU AUA CAC AGG CGG GGC AUU 10025
 Ile Arg His Ala Leu Ile Ser Ser Ile Ile His Arg Arg Gly Ile
 GUG AAA GAA UAU UUG UUC CAG UAU UGC CAG GGU AUG CCA UCA GGG 10070
 Val Lys Asp Tyr Leu Phe Gln Tyr Cys Gln Gly Met Pro Ser Gly
 20 UUC GCC AUG ACA GUG AUA UUC AAU UCG UUU AUG AAC UAU UAU UAU 10115
 Phe Ala Met Thr Val Ile Phe Asn Ser Phe Met Asn Tyr Tyr Tyr
 CUG UCU UUG GCC UGG AUG AAU CUG AUA AGU GCA UCC CCC CUU AGU 10160
 Leu Ser Leu Ala Trp Met Asn Leu Ile Ser Ala Ser Pro Leu Ser
 CCA CAA GCU UCU UUG AGA UAU UUU GAU GAG UAU UGU AAG GUC AUU 10205
 25 Pro Gln Ala Ser Leu Arg Tyr Phe Asp Glu Tyr Cys Lys Val Ile
 GUU UAC GGU GAU GAU AAU AUU GUU GCC GUC AAC GAA GAA UUC UUA 10250
 Val Tyr Gly Asp Asp Asn Ile Val Ala Val Asn Asp Asp Phe Leu
 GAG UAC UAU AAC UUG AGG CUU GUG GCA GGC UAU CUU AGU CAA UUU 10295
 Glu Tyr Tyr Asn Leu Arg Leu Val Ala Gly Tyr Leu Ser Gln Phe
 30 GGA GUA AGC UAC ACU GAU GAC GCC AAG AAC CCA AUA GAG AAG AGC 10340
 Gly Val Ser Tyr Thr Asp Asp Ala Lys Asn Pro Ile Glu Lys Ser
 GAA CGA UAU GUG AAG AUA GAA GAC GUU ACG UUC UUA AAA CGG CGA 10385
 Asp Arg Tyr Val Lys Ile Asp Asp Val Thr Phe Leu Lys Arg-Arg
 UGG GUG AGU CUU GGC GGU AGA GCU UCG AUG CUG UAC AAA GCU CCG 10430
 35 Trp Val Ser Leu Gly Gly Arg Ala Ser Met Leu Tyr Lys Ala Pro

	CUU GAC AAG GUU AGC AUU GAG GAA AGG CUU AAC UGG AUC AGA GAG	10475
	Leu Asp Lys Val Ser Ile Glu Asp Arg Leu Asn Trp Ile Arg Glu	
	UGU GAC GAU GGG GAA CUA GCU CUG GUG CAG AAC AUU GAA AGU GCU	10520
	Cys Asp Asp Gly Asp Leu Ala Leu Val Gln Asn Ile Asp Ser Ala	
5	CUG UAC GAA GCU AGU AUU CAU GGC CAC ACA UAU UUU GGA GAG CUU	10565
	Leu Tyr Asp Ala Ser Ile His Gly His Thr Tyr Phe Gly Glu Leu	
	AAA GAU AAA AUU GCU AAA GCC UGU GAU GCA GUC AUG AUA ACU AUG	10610
	Lys Asp Lys Ile Ala Lys Ala Cys Asp Ala Val Met Ile Thr Met	
	CCA AAU AUA AGA UAU AUU GAC UGC CAG AGA CGA UGG UGG ACC UCC	10655
10	Pro Asn Ile Arg Tyr Ile Asp Cys Gln Arg Arg Trp Trp Thr Ser	
	AUG ACU GGU GGG UAU CUU GAG CCG UCU GAU GUC ACC AAA CUU GUA	10700
	Met Thr Gly Gly Tyr Leu Glu Pro Ser Asp Val Thr Lys Leu Val	
	AGG CUU GUU GAG AAA GGA CUA CUA GAC CCG AAA UCA GUA UGG AAA	10745
	Arg Leu Val Glu Lys Gly Leu Leu Asp Pro Lys Ser Val Trp Lys	
15	GAC CCA UUG UAC AGA ACC AAC AAG UUG CUA UUC GAC CUA UUG AGG	10790
	Asp Pro Leu Tyr Arg Thr Asn Lys Leu Leu Phe Asp Leu Leu Arg	
	GAG GUU AAG GCA GCA CCC CUG GCC GCA UUU GUG GUC UAA	10829
	Glu Val Lys Ala Ala Pro Leu Ala Ala Phe Val Val Stop	
	GUUACCCUUC UGACAAAAGG GCCUUGAACG GUUAUGGUUG AACAGAACUG	10879
20	UAAAAGGUGA GGACUAUUAU AGUUGUAGUA CGGAUGAGAU UGAAAGAAAA	10929
	UUGGGUCACU CCCAUUCCUU UAUUAGGAAG GAGUGAUACC UUUUGUGUAG	10979
	AUCUCUACCC CGAAACUCUU GAACCCUCAC ACGUUUUGGA GUAACCAGUA	11029
	CACCCUUUUA GGUGGACCCU CGACUAUAGA UCGAGACCAA GUAUUGACUU	11079
	GGUGUUCACG UCUUGCCGGA CGCAAAUUGG CACCCUUGUU UAGUGAUUUC	11129
25	AAGGUUACAA AUGUCACGCC CCACUAGUAA AAGUUUUGGU AUUAACGCAU	11179
	UCGAACCGCC AAUGUAUACG UGUUUUCCCU UUUACUUUUU GUAUGUCGUC	11229
	GUGGUGACGA GAUGCACGCC UGGUCAGCGG GGAUAAGUU CACUAUAUGA	11279
	ACAGACUCCG GCGAGCGAGA CACGCUGUCG GCCUCGGGAG AGGGAACUAG	11329
	CUCCAGGCAC UUAUUUCCUG AAGUGUUAGA ACUAAGCGUU UGAUCCUCCU	11379
30	CCGGGGGAAA GAGAACGCCA GUUCUUUAAG CCAUAACUCU AGUGAGUUGA	11429
	AUCCUAUUCA UCCUUCUUAU GAUUAAGGAU UUCUGAAGUC UAUCAUGAAA	11479
	AGUAGAUAGA AAGCAACACG UCAAUAACGU GGAACCUUUU CCGAGGAAGU	11529
	AGGGUGCUUG UUCGAAAAUC AUGGUAGAUU CGGAAACAAU UUGC UUAGAG	11579
	UGUGUCUUUU CGCGUUGGUA GUUCAACCGU UAGGGCUAGG CACACUUCUC	11629
35	CACGGGUUUG UGCUGCAGUA UUAUAUAUCA UUAAGGUACU GUGCUAUAGC	11679

GGAGAAAUUA CAAAGCGUUG AACACAUUGA CGAUGGGGCC CAAUGCGCAC 11729
 CCGGAUGUGU UACGCACCGU UUUUCUCUGU GUCACUAUAG AUAAAAGUGG 11779
 GGUAGC-polyA 11785

(2) INFORMATION FOR SEQ ID NO: 5:

5 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 bases
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: viral RNA

(A) DESCRIPTION: RNA codons for first 15 amino acids at 5' end of MCDV coat protein 1 (CP1)

(iii) HYPOTHETICAL: No

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

15 GUU UCA UUG GGU CGG UCA UUU GAG AAU GGA GUG CUU AUU GGU AGU 45
 Val Ser Leu Gly Arg Ser Phe Glu Asn Gly Val Leu Ile Gly Ser

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(A) DESCRIPTION: first 15 amino acids of MCDV coat protein 3

25 (iii) HYPOTHETICAL: No

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Leu Gln Val Ala Ser Leu Thr Asp Ile Gly Asp Leu Ser Ser Val 15

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

35 (A) DESCRIPTION: first 15 amino acids of MCDV coat protein 1

(iii) HYPOTHETICAL: No

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Val Ser Leu Gly Arg Ser Phe Glu Asn Gly Val Leu Ile Gly Ser 15

WHAT IS CLAIMED IS:

1. A DNA clone coding substantially solely for a coat protein of maize dwarf mosaic virus.
- 5 2. An expression cassette comprising a DNA clone according to Claim 1, operably linked to plant regulatory sequences which cause the expression of the DNA clone in plant cells.
3. An expression cassette comprising a DNA clone according to Claim 1, operably linked to bacterial expression regulatory sequences which cause the
10 expression of the DNA clone in bacterial cells.
4. Bacterial cells containing as a foreign plasmid at least one copy of an expression cassette according to Claim 3.
5. Transformed plant cells containing as foreign DNA at least one copy of the DNA sequence of an expression cassette according to Claim 2.
- 15 6. Transformed cells according to Claim 5, further characterized in being cells of a monocotyledonous species.
7. Transformed cells according to Claim 6, further characterized in being maize, sorghum, wheat or rice cells.
8. Transformed cells according to Claim 5, further characterized in being
20 cells of a dicotyledonous species.
9. Transformed cells according to Claim 8, further characterized in being soybean, alfalfa, tobacco or tomato cells.
10. A maize cell or tissue culture comprising cells according to claim 7.
11. A transformed maize plant, the cells of which contain as foreign DNA
25 at least one copy of the DNA sequence of an expression cassette according to Claim 2.
12. A method of imparting resistance to maize chlorotic dwarf virus and maize dwarf mosaic virus - A to plants of a MCDV or MDMV-A susceptible taxon, comprising the steps of:
30 a) culturing cells or tissues from at least one plant from the taxon,
b) introducing into the cells of the cell culture or tissue culture at least one copy of an expression cassette comprising a DNA clone from the RNA genome of MCDV which codes substantially solely for the coat protein of the virus, operably linked to plant regulatory sequences which cause the
35 expression of the DNA clone in the cells, and

c) regenerating MCDV-resistant whole plants from the cell culture or tissue culture.

13. A method according to Claim 12 which comprises the further step of sexually or clonally reproducing the whole plants in such manner that at least one
5 copy of the sequence provided by the expression cassette is present in the cells of progeny of the reproduction.

14. A method according to Claim 12 in which the expression cassette is introduced into the cells by electroporation.

15. A method according to Claim 12 in which the expression cassette is introduced into the cells by microparticle bombardment.

16. A method according to Claim 12 in which the expression cassette is introduced into the cells by microinjection.

17. A method according to Claim 13 for providing MCDV and MDMV-A resistance in *Agrobacterium tumefaciens*-susceptible dicotyledonous plants in which
15 the expression cassette is introduced into the cells by infecting the cells with *Agrobacterium tumefaciens*, a plasmid of which has been modified to include the expression cassette.

18. A method of imparting resistance to maize chlorotic dwarf virus and maize dwarf mosaic virus strain A to plants of a MCDV or MDMV-A susceptible
20 taxon, comprising the steps of:

a) selecting a fertile, MCDV resistant plant prepared by the method of Claim 12 from a sexually compatible taxon;

b) sexually crossing the MCDV resistant plant with a plant from the MCDV susceptible taxon;

25 c) recovering reproductive material from the progeny of the cross; and

d) growing resistant plants from the reproductive material.

19. A method according to Claim 18 which comprises the further steps of repetitively:

30 a) backcrossing the MCDV resistant progeny with MCDV susceptible plants from the susceptible taxon; and

b) selecting for expression of MCDV resistance among the progeny of the backcross,

until the desired percentage of the characteristics of the susceptible taxon are present in the progeny along with MCDV resistance.

20. A DNA molecule coding for maize chlorotic dwarf virus or a portion thereof which is capable of conferring resistance to maize chlorotic dwarf virus when expressed in a plant cell.

Figure 1

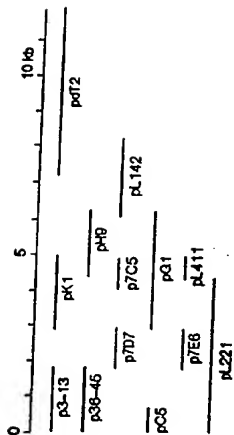
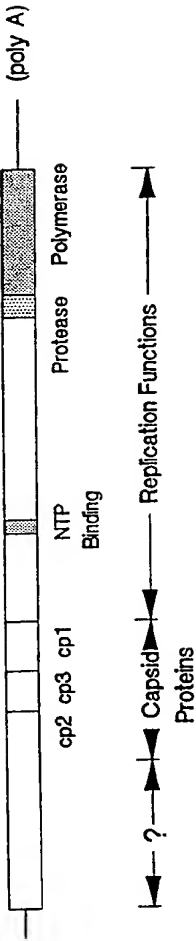


Figure 2





INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12N 15/40, 15/82, 5/10, A01H 5/00

A3

(11) International Publication Number: WO 94/21796

(43) International Publication Date: 29 September 1994 (29.09.94)

(21) International Application Number: PCT/US94/03028

(22) International Filing Date: 22 March 1994 (22.03.94)

(30) Priority Data:
08/038,768 24 March 1993 (24.03.93) US

(71) Applicants: PIONEER HI-BRED INTERNATIONAL, INC.
[US/US]; 700 Capital Square, 400 Locust Street, Des
Moines, IA 50309 (US). THE UNITED STATES OF
AMERICA as represented by THE SECRETARY, UNITED
STATES DEPARTMENT OF AGRICULTURE [US/US];
12th and Independence Avenue, S.W., Washington, DC
20250-1400 (US).

(72) Inventors: ROTH, Bradley, A.; 210 N.W. 3rd Place, Grimes,
IA 50111 (US). TOWNSEND, Rod; 541 Waterbury Circle,
Des Moines, IA 50312 (US). MCMULLEN, Michael, D.;
1680 Madison Avenue, Wooster, OH 44691-4096 (US).

(74) Agents: ROTH, Michael, J. et al.; 700 Capital Square, 400
Locust Street, Des Moines, IA 50309 (US).

(81) Designated States: AT, AU, BB, BG, BR, BY, CA, CH, CN,
CZ, DE, DK, ES, FI, GB, HU, JP, KP, KR, KZ, LK, LU,
LV, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD,
SE, SK, UA, UZ, VN, European patent (AT, BE, CH, DE,
DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI
patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE,
SN, TD, TG).

Published

With international search report.

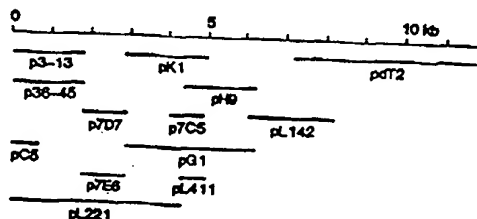
Before the expiration of the time limit for amending the
claims and to be republished in the event of the receipt of
amendments.

(88) Date of publication of the International search report:
10 November 1994 (10.11.94)

(54) Title: MAIZE CHLOROTIC DWARF VIRUS AND RESISTANCE THERETO

(57) Abstract

Methods and materials are provided to isolate the coat protein genes from maize chlorotic dwarf virus. One or more of these genes (MCDV-CP₁, MCDV-CP₂ or MCDV-CP₃) is then incorporated in an expression cassette designed for suitable expression in a plant cell system. The resulting transformation vector is then introduced into maize to provide cross protection to MCDV or related viral infections.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 94/03028

A. CLASSIFICATION OF SUBJECT MATTER

IPC 5 C12N15/40 C12N15/82 C12N5/10 A01H5/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 5 C12N A01H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	PHYTOPATHOLOGY vol. 78 , 1988 , 12 PART 1 page 1599 JILKA, J., ET AL. 'Cloning and sequencing of the coat protein cistron of maize dwarf mosaic virus, strains A and B'	1
Y	see abstract 689 --- -/--	2,5-11

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understate the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *A* document member of the same patent family

Date of the actual completion of the international search

31 August 1994

Date of mailing of the international search report

21 -09- 1994

Name and mailing address of the ISA

European Patent Office, P.B. 5118 Patentaan 2
NL - 2280 HV Rijswijk
Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,
Fax (+ 31-70) 340-3016

Authorized officer

Maddox, A

INTERNATIONAL SEARCH REPORT

 Internat'l Application No
 PCT/US 94/03028

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Quotation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CHEMICAL ABSTRACTS, vol. 115, no. 11, 1991, Columbus, Ohio, US; abstract no. 108765, FRENKEL, M.J., ET AL. 'Unexpected sequence diversity in the amino-terminal ends of the coat proteins of strains of sugarcane mosaic virus' see abstract & J. GEN. VIROL.	1 2,5-11
Y	vol. 72, no. 2, 1991 pages 237 - 242 ---	2,5-11
Y	EP,A,0 223 452 (MONSANTO) 27 May 1987 see claim 13 ---	2,5-11
X	DATABASE BIOSIS BIOSCIENCES INFORMATION SERVICE, PHILADELPHIA, PA, US ABSTRACT BR45:70542 BERNARDI, F., ET AL. 'Cloning the N-terminal region of maize dwarf mosaic virus coat gene into bacteria and characterizing the expression product' see abstract & PHYTOPARASITICA. 14TH CONGRESS OF THE ISRAELI PHYTOPATHOLOGICAL SOCIETY, FEB 15-16, 1993. vol. 21, no. 2, 1993 page 154 ---	1,3,4
X	CHEMICAL ABSTRACTS, vol. 115, no. 11, 1991, Columbus, Ohio, US; abstract no. 107314, JILKA, J.M. 'Cloning and characterization of the 3' terminal regions of RNA from select strains of maize dwarf mosaic virus and sugarcane mosaic virus' page 210 ; see abstract & PHD THESIS, UNIVERSITY OF ILLINOIS, URBANA, ILLINOIS, USA. 1990 & DISS. ABSTR. INT. B. vol. 51, 1991, 12 PART 1 page 5719 ---	1
X	DATABASE BIOSIS BIOSCIENCES INFORMATION SERVICE, PHILADELPHIA, PA, US BR38:57628 GE, X., ET AL. 'Characterization of maize chlorotic dwarf virus MCDV rna' see abstract & PHYTOPATHOLOGY vol. 79, no. 10, 1989 page 1157 ---	20
1 A	---	12-19 -/--

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

page 2 of 3

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO,A,93 14210 (SANDOZ) 22 July 1993 see the whole document ---	1,2,5-11
A	DATABASE BIOSIS BIOSCIENCES INFORMATION SERVICE, PHILADELPHIA, PA, US BR38:57625 MAROON, C.M., ET AL. 'Serological relationships of the capsid proteins of the type isolate of maize chlorotic dwarf virus MCDV-T' see abstract & PHYTOPATHOLOGY vol. 79, no. 10, 1989 page 1157 ---	1-20
A	CHEMICAL ABSTRACTS, vol. 115, no. 7, 1991, Columbus, Ohio, US; abstract no. 65640, GE, X. 'Characterization of the genome of maize chlorotic dwarf virus and an associated satellite virus' see abstract & Dissertation available, univ. microfilms int., order no. DA 9105112 Diss. abstr. int. B 51(10), 4666. ---	1-20
A	DATABASE BIOSIS BIOSCIENCES INFORMATION SERVICE, PHILADELPHIA, PA, US BR46:293 KOONIN, E.V., ET AL. 'Evolution and taxonomy of positive-strand RNA viruses: Implications of comparative analysis of amino acid sequence' see abstract & CRITICAL REVIEWS IN BIOCHEMISTRY AND MOLECULAR BIOLOGY vol. 28, no. 5, 1993 pages 375 - 430 ---	1-20
A	BIOLOGICAL ABSTRACTS, vol. 63 Philadelphia, PA, US; abstract no. 6035, GINGERY, R.E. 'Properties of maize chlorotic dwarf virus and its ribonucleic acid' see abstract & VIROLOGY vol. 73, no. 2, 1976 pages 311 - 318 -----	1-20

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 94/03028

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A-0223452	27-05-87	AU-B- 608204 AU-A- 6452886 JP-A- 62201527	28-03-91 30-04-87 05-09-87
WO-A-9314210	22-07-93	NONE	

Form PCT/ISA/210 (patent family annex) (July 1992)